

## **Rodrigo Moreira Garcia**

**A busca da informação especializada e a efetividade  
de sua recuperação:**  
interação entre bibliotecário, usuário e base de dados



Marília – SP  
2005

**RODRIGO MOREIRA GARCIA**

**A busca da informação especializada e a efetividade  
de sua recuperação:  
interação entre bibliotecário, usuário e bases de dados**

Trabalho de Conclusão de Curso apresentado ao Conselho de Curso de Biblioteconomia da Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, como requisito parcial para a obtenção do título de Bacharel em Biblioteconomia.

**Linha de pesquisa:** Informação e Sociedade.

**Orientadora:** Dr<sup>a</sup> Helen de Castro Silva

**Bolsa de Iniciação Científica:** *Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP.*

Marília – SP  
2005

Garcia, Rodrigo Moreira  
G216b A busca da informação especializada e a efetividade de sua  
recuperação: interação entre bibliotecário, usuário e base de dados  
/ Rodrigo Moreira Garcia. – Marília: UNESP, 2005  
189 f: il.; 30 cm

Trabalho de Conclusão de Curso - **TCC** (Graduação) –  
Faculdade de Filosofia e Ciências – Universidade Estadual  
Paulista, Marília, 2005.

1 Recuperação da informação. 2 Comportamento de busca.  
3 Estratégia de busca. 4 Base de dados. 5 Informação  
especializada. 6 Usuário final. I. Autor II. Título



CDD 025.5

# **RODRIGO MOREIRA GARCIA**

## **A busca da informação especializada e a efetividade de sua recuperação:**

interação entre bibliotecário, usuário e base de dados

Trabalho de Conclusão de Curso apresentado ao Conselho de Curso de Biblioteconomia da Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, como requisito parcial para a obtenção do título de Bacharel em Biblioteconomia.

**Linha de pesquisa:** Informação e Sociedade.

**Bolsa de Iniciação Científica:** Fundação de Amparo à Pesquisa do Estado de São Paulo – **FAPESP**.

Data de defesa : 21/01/2005

### **BANCA EXAMINADORA:**

---

**Nome:** Helen de Castro Silva

**Titulação:** Prof<sup>a</sup>. do Departamento de Ciência da Informação, Faculdade de Filosofia e Ciências, UNESP – Campus de Marília.

---

**Nome:** Silvana Aparecida Borseti Gregório Vidotti

**Titulação:** Prof<sup>a</sup>. do Departamento de Ciência da Informação, Faculdade de Filosofia e Ciências, UNESP – Campus de Marília.

---

**Nome:** Maria Luzinete Euclides

**Titulação:** Bibliotecária, Faculdade de Filosofia e Ciências  
UNESP – Campus de Marília.

**Local:** Universidade Estadual Paulista – UNESP  
Faculdade de Filosofia e Ciências  
**Campus de Marília**

## DEDICATÓRIA

Chegar até aqui é, sem sombra de dúvida, uma grande conquista. É neste sentido que dedico este trabalho às pessoas mais importantes e que sempre estiveram e estarão presentes em minha vida.

Ao meu pai e grande herói da minha vida, João Roberto Garcia. Obrigado pelo grande amor, compreensão e dedicação;

À minha mãe, Edite Maria Moreira Garcia, sempre companheira e dedicada. Obrigado pelo amor, compreensão e dedicação;

Ao meu irmão e futuro Jornalista, Ricardo Moreira Garcia, pelo amor e amizade de sangue. Que minha trajetória até aqui sirva de exemplo;

Aos meus tios (tia Lair & tio Vicente, e a tia Sônia), pelo carinho e ajuda nos momentos difíceis;

Às minhas primas Simone e Silene, pelo carinho e amizade, apesar do pouco convívio que tivemos;

Aos meus priminhos Gabriel e Bárbara, pela alegria e vida que vocês têm dado a todos da família;

A todos os meus familiares de Santa Rosa do Viterbo, São Paulo e *España*;

Aos meus futuros sogros Maria Mercês Flamino e José Magalhães Guedes e família, pelo apoio e carinho (apesar da distância);

**E a você, meu grande e eterno amor, Adriana Nascimento Flamino, que sempre estive ao meu lado em todos os momentos, desde que estamos juntos, sendo companheira, amiga, namorada, noiva, e brevemente minha linda futura esposa. "Onde quer que eu vá, o que eu faça, sem você, não tem graça". Sem você nada disso teria sentido. Amo-te, minha Dri.**

**A esta criaturinha que veio ao mundo um pouquinho apressado mas, ao mesmo tempo, no momento certo, fruto do grande amor entre mim e a mãe dele, Adriana Nascimento Flamino, e que só nos vem trazendo alegria e felicidade. Que esta alegria e vontade de viver, meu amado filho Ian Pablo Flamino Garcia, lhe acompanhem por toda a sua jornada nesta Terra.**

## AGRADECIMENTOS

Agradeço a todos que tornaram possível a realização deste trabalho e, em especial:

À Fundação de Amparo a Pesquisa do Estado de São Paulo - **FAPESP**, pelo apoio financeiro para o desenvolvimento desta pesquisa de Iniciação Científica;

À Prof<sup>a</sup> Helen de Castro Silva, pelo apoio, estímulo e dedicada orientação;

À professora Silvana Aparecida Borseti Gregório Vidotti, por fazer parte de minha banca examinadora e pelas pertinentes sugestões dadas ao meu trabalho de conclusão de curso;

À bibliotecária Maria Luzinete Euclides, pelas contribuições e por fazer parte de minha banca examinadora e pelas pertinentes sugestões dadas ao meu trabalho de conclusão de curso;

À professora e Consultora Sylvia Helena Morales Horiguela de Moraes (Escritório de Pesquisa da UNESP) e a "Marizinha" (Kimiko Marisa Takahashi), pela atenção e disposição para comigo.

À Vânia M. S. Reis Fantin (Diretora da Biblioteca do campus de Marília), Hiromi, Célia, Tina (Bibliotecárias), Zilda, Sônia, Lair, Ilma, Akiko, Neuza e Telma (funcionários), e estagiários pelo apoio, colaboração, amizade e compreensão sempre.

Aos alunos da Pós-Graduação que contribuíram com esta pesquisa;

À D. Cida e família por cuidarem do Ian Pablo nos períodos de aula e pela amizade;

Aos funcionários do Res. André Luís: Sr. Oswaldo, José e Sr. Luís, pela amizade e dedicação;



Aos meus amigos de São Paulo, Minelli (e família) e Carlos Jr. e família ("[The Minduins Forever](#)"), pela longa amizade; Duda (e família), Rick, Willian, Barbosa (e família), Lubi, Renatinho, Gordo (Primo do Carlos), Mariusz (polonês), Marcelo (ponês), Carlão (onde estiver), Fernando "Aroldo" (onde estiver), Abrão (onde estiver), Yunes (onde estiver), pelos bons momentos que todos um dia me proporcionaram.

Agradeço a todos aqueles que durante esta peregrinação pelo deserto, estenderam a mão e foram verdadeiramente **amigos**.

A cega arrogância que vem com a juventude é paralisada pela falta de flexibilidade do velho; a impaciência e a ambição do jovem chocam-se com o medo que o velho tem de idéias novas e de suas possíveis conseqüências contra a sua hegemonia. A curto prazo o velho em geral vence, o jovem recuando para repensar seu plano de ataque. Se a força dos argumentos do jovem, contudo, for realmente grande, ele conquistará o velho no final, forçando uma completa transformação de valores ou, pelo menos, o início de uma transição (GLEISER<sup>α</sup>, 1997).

É muito tentador pararmos para escutar os elogios embevecedores. Mas a única coisa a ser feita é dar as costas a isso e continuar a trabalhar. Trabalho. Não existe mais nada.  
Albert Einstein

A estrela então explode com uma fúria tremenda, num fenômeno conhecido como explosão do tipo *supernova*. Portanto, o carbono, o oxigênio e outros elementos pesados, que não só fazem parte de nosso organismo como também são fundamentais para nossa sobrevivência, foram sintetizados no interior de estrelas moribundas antes de serem projetados através do espaço interestelar. Nós somos filhos das estrelas (GLEISER<sup>α</sup>, 1997, p. 377).

[...] mesmo que horizontes possam existir, eles são horizontes em fuga, que nunca serão atingidos; numa terra de horizontes em fuga, um viajante inspirado sempre encontrará novas maravilhas. Pelo menos, essa é a minha metáfora para a criatividade humana (GLEISER<sup>α</sup>, 1997, p. 360).

---

<sup>α</sup> GLEISER, Marcelo. **A dança do universo**: dos mitos de criação ao Big Bang. São Paulo: Companhia das Letras, 1997.

GARCIA, Rodrigo Moreira. **A busca da informação especializada e a efetividade de sua recuperação**: interação entre bibliotecário, usuário final e base de dados. 2005. 189f. Trabalho de Conclusão de Curso – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, Marília. 2005.

## RESUMO

As atividades de busca e recuperação da informação consistem em selecionar da massa informacional disponível, aqueles elementos que atenderem satisfatoriamente as necessidades informacionais de um grupo ou indivíduo, no instante de sua solicitação. Essas atividades, são próprias do profissional bibliotecário, visto que, a essência da biblioteconomia consiste em localizar, tratar, armazenar, recuperar e disseminar a informação. Mas, com o desenvolvimento das tecnologias de informação e comunicação a capacidade de armazenamento em bases de dados vem crescendo, assim como o acesso direto dos usuários à informação, tornando a recuperação da informação em meio eletrônico uma tarefa bastante complexa. Estes fatores levam a indagar se os usuários têm conseguido suprir suas necessidades informacionais eficientemente, sem a intermediação do profissional bibliotecário. Foi proposto o estudo do comportamento de uma amostra dos alunos da Pós-Graduação da UNESP-Marília nas atividades de busca e recuperação de informação em bases de dados especializadas, através da aplicação de questionários e observações de seus procedimentos de busca. Com apoio de referencial teórico pertinente ao tema, os dados foram tabulados e analisados quantitativa e qualitativamente. Como resultados, constatou-se que diversos problemas impedem que os usuários utilizem-se de toda a potencialidade que as bases de dados atualmente oferecem para a recuperação da informação. Conclui-se que, os estudos sobre o comportamento de busca dos usuários e os conhecimentos inerentes ao bibliotecário, podem contribuir para o design de sistemas de recuperação da informação e interfaces de busca mais eficientes.

**Palavras-chave:** Base de dados; Bibliotecário; Comportamento de Busca; Estratégia de busca; Informação especializada; Recuperação da informação; Usuário final.



GARCIA, Rodrigo Moreira. **A busca da informação especializada e a efetividade de sua recuperação**: interação entre bibliotecário, usuário final e base de dados. 2005. 189f. Trabalho de Conclusão de Curso – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista – UNESP, Marília. 2005.

## **ABSTRACT**

The search activities and information retrieval consist of selecting of the information mass those elements that It assist satisfactorily the information needs of a group or individual, in the instant of your solicitation. Those activities, are own of the professional librarian, because, the essence of the Librarianship consists of locating, to treat, to store, to retrieval and to disseminate the information. But, with the development of the information and communication technologies the storage capacity in databases is growing, as well as the direct access of the users to the information, turning the information retrieval in half electronic a quite complex task. These factors take to investigate if the users have been getting to supply your information needs efficiently, without the professional librarian as intermediate. It was proposed the study of the behavior of a sample of the students of the powders graduation of the UNESP-Marília in the search activities and information retrieval in databases specialized, through the application of questionnaires and observations of your search procedures. With support of pertinent theoretical referential to the theme, the data were tabulated and analyzed quantitative and qualitative. As results, it was verified that several problems impede that the users take advantage of all the potentiality that the bases of data now offer for the information retrieval. It is ended that, the studies on the behavior of the users' search and the inherent knowledge to the librarian, they can contribute to the design of systems of information retrieval and more efficient search interfaces.

**Keywords:** Database; Librarian; Search Behavior; Search Strategy; Information Specialized; Information Retrieval; End-User.

## LISTA DE ILUSTRAÇÕES

<b>FIGURA 1:</b> Arquitetura genérica dos agentes de filtragem de informação	33
<b>FIGURA 2:</b> Ambiente e seus principais componentes	35
<b>FIGURA 3:</b> Gerações de sistemas de recuperação da informação	44
<b>FIGURA 4:</b> Vocabulário Controlado: vantagens e desvantagens	55
<b>FIGURA 5:</b> Linguagem Natural: vantagens e desvantagens	56
<b>FIGURA 6:</b> Exemplo de fluxo da representação do item documentário para a recuperação da informação	60
<b>FIGURA 7:</b> Esquema dos componentes do modelo interativo	72
<b>FIGURA 8:</b> Processo de interação da Recuperação da Informação	74
<b>FIGURA 9:</b> Processo da Recuperação da Informação	75
<b>FIGURA 10:</b> Resolução do problema no processo de busca de informação	76
<b>FIGURA 11:</b> Recursos de busca usuais	82
<b>FIGURA 12:</b> Operadores lógicos booleanos	83
<b>FIGURA 13:</b> Modelo metodológico de pesquisa de análise teórico-empírica	106
<b>FIGURA 14:</b> Portal dos Sistemas de Bibliotecas das Universidades Estaduais Paulistas – CRUESP Bibliotecas	114
<b>FIGURA 15:</b> UnibibliWeb	115
<b>FIGURA 16:</b> Biblioteca Eletrônica do CRUESP	116
<b>FIGURA 17:</b> Portal Bibliotecas UNESP	117
<b>FIGURA 18:</b> Portal de Periódicos da CAPES	118
<b>FIGURA 19:</b> Portal da Pesquisa	119
<b>FIGURA 20:</b> Modelo de interação na recuperação da informação (adaptado)	163

## LISTA DE GRÁFICOS

<b>GRÁFICO 1:</b> Número de sujeitos por amostragem estratificada proporcional para aplicação de questionário _____	110
<b>GRÁFICO 2:</b> Número de participantes por amostragem estratificada proporcional para realização de observação _____	112
<b>GRÁFICO 3:</b> Formação acadêmica dos alunos sujeitos da pesquisa _____	123
<b>GRÁFICO 4:</b> Profissão/ocupação dos alunos sujeitos da pesquisa _____	124
<b>GRÁFICO 5:</b> Opinião dos usuários sobre a infra-estrutura oferecida _____	126
<b>GRÁFICO 6:</b> Dificuldades de busca em bases de dados especializadas _____	128
<b>GRÁFICO 7:</b> Quem realiza levantamentos bibliográficos _____	129
<b>GRÁFICO 8:</b> Participação em curso ou treinamento em bases de dados especializadas_	135
<b>GRÁFICO 9:</b> Representação gráfica das fontes utilizadas na busca e recuperação de informações _____	138
<b>GRÁFICO 10:</b> Planejamento da estratégia de busca _____	141
<b>GRÁFICO 11:</b> Satisfação dos usuários quanto as próprias buscas _____	143
<b>GRÁFICO 12:</b> Familiaridade com os instrumentos de auxílio à busca _____	144
<b>GRÁFICO 13:</b> Utilização dos Instrumentos de auxilio a busca em bases de dados _____	145

## LISTA DE TABELAS

<b>TABELA 1:</b> Número de alunos matriculados nos programas de Pós-Graduação da Unesp Marília em 2004 _____	109
<b>TABELA 2:</b> Local onde o usuário costuma realizar buscas em bases de dados _____	124
<b>TABELA 3:</b> Motivos para a procura pelo bibliotecário ou outros para a realização de levantamentos bibliográficos _____	130
<b>TABELA 4:</b> Dificuldades encontradas pelos usuários em levantamentos bibliográficos em Bases de dados especializadas _____	132
<b>TABELA 5:</b> Motivos para solicitações de auxílio ao bibliotecário de referência _____	134
<b>TABELA 6:</b> Contribuições de curso ou treinamento para o uso de bases de dados especializadas _____	136
<b>TABELA 7:</b> Bases de dados mais utilizadas, segundo nº de usuários _____	140
<b>TABELA 8:</b> Procedimentos adotados pelos usuários quando os resultados são insatisfatórios na busca em base de dados _____	142
<b>TABELA 9:</b> Nº de registros relevantes recuperados por nº de registros recuperados _____	161

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
<b>2</b>	<b>RECUPERAÇÃO DA INFORMAÇÃO</b>	<b>19</b>
2.1	Mas o que seria Informação ?	24
2.2	Por que Recuperar a Informação ?	27
2.3	Recuperação da Informação e as Tecnologias da Informação e Comunicação	29
2.4	As Bases de Dados	37
<b>3</b>	<b>RECUPERAÇÃO DA INFORMAÇÃO EM BASES DE DADOS</b>	<b>47</b>
3.1	Representação Descritiva ou Catalogação	48
3.2	Representação temática ou Indexação e as Linguagens de indexação	49
3.3	Metadados	59
3.3.1	Formato de Intercâmbio MARC 21	63
3.3.2	EXtensible Markup Language – XML	64
3.4	Armazenamento	68
3.5	Etapas do Processo de Recuperação da Informação	69
3.6	Estratégias de busca	76
<b>4</b>	<b>SERVIÇO DE REFERÊNCIA E A RECUPERAÇÃO DA INFORMAÇÃO PELO USUÁRIO FINAL</b>	<b>88</b>
4.1	Abordagens na Recuperação da Informação	102
<b>5</b>	<b>PROCEDIMENTOS METODOLÓGICOS</b>	<b>106</b>
<b>6</b>	<b>APRESENTAÇÃO E ANÁLISE DOS RESULTADOS</b>	<b>113</b>
6.1	Infra-estrutura disponível para pesquisa bibliográfica	113
6.2	Necessidade informacional, conhecimentos e dificuldades na utilização das bases de dados pelos usuários	123
6.3	Necessidade de otimização dos processos de planejamento e operacionalização da estratégia de busca em bases de dados	147
<b>7</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>164</b>
	<b>REFERÊNCIAS</b>	<b>170</b>
	<b>BIBLIOGRAFIA CONSULTADA</b>	<b>179</b>
	<b>APÊNDICE</b>	<b>183</b>

## CAPÍTULO 1

O saber, o conhecimento, é considerado cumulativo e, para muitos, infinito.  
FIGUEIREDO<sup>1</sup>

### INTRODUÇÃO

O presente estudo insere-se na linha de pesquisa Informação e Sociedade, do Departamento de Ciência da Informação da Universidade Estadual Paulista – UNESP, Faculdade de Filosofia e Ciências – FFC, Campus de Marília, tendo como tema a **Recuperação da Informação**, especificamente, o comportamento de usuários na busca da informação especializada em bases de dados. Trata-se de Trabalho de Conclusão de Curso – TCC, com bolsa de **Iniciação Científica**, concedida pela Fundação de Amparo à Pesquisa do Estado de São Paulo – **FAPESP**, após apreciação de projeto, sob o processo nº 03/14144-2, referente ao período de março a dezembro de 2004.

Por atualmente estarmos vivendo no liminar da era da informação, na qual, recuperar a informação necessária para as tomadas de decisões torna-se imprescindível para a sobrevivência das organizações, sejam elas públicas ou privadas, destinadas ao capital ou à pesquisa.

As atividades de busca e recuperação da informação consistem em selecionar da massa informacional disponível, aqueles elementos que atenderem satisfatoriamente as necessidades ou interesses de informação de um grupo ou indivíduo, no instante de sua solicitação. Considerando que a essência das atividades biblioteconômicas consiste em localizar, tratar, armazenar, recuperar e disseminar a informação é certo, então, que o

---

<sup>1</sup> FIGUEIREDO, Nice Menezes de. **Paradigmas modernos da Ciência da informação**. São Paulo: Polis, 1999, 168 p.

profissional bibliotecário é o intermediário que dispõe de maior habilidade durante o processo de planejamento e operacionalização da estratégia de busca para uma efetiva busca e recuperação da informação. Independente dos tipos de formatos ou suportes onde esta informação se encontre, ele tem condições de oferecer a seus usuários e/ou clientes os subsídios necessários para a resolução de seus problemas ou questões, no tempo mais hábil possível.

Mas visto que, cada vez mais, os avanços das tecnologias de informação e de comunicação ampliam as facilidades de acesso à informação pelos usuários finais levanta-se a questão: os clientes/usuários têm conseguido suprir suas necessidades informacionais eficientemente, sem a intermediação do profissional bibliotecário? A hipótese levantada neste trabalho para esta questão é a de que o cliente/usuário, é capaz de recuperar a informação necessária com efetividade, desde que domine os processos de planejamento e operacionalização das estratégias de busca para uma efetiva recuperação da informação.

O tema desta pesquisa sobre o comportamento e desempenho dos usuários finais na busca e recuperação da informação em bases de dados especializadas, é atual e de extrema relevância, visto que o desenvolvimento das tecnologias de informação e comunicação vem aumentando a capacidade de armazenamento em bases de dados, assim como tem facilitado o acesso e a autonomia de clientes/usuários, tornando os processos de busca e recuperação da informação em meio eletrônico uma tarefa bastante complexa. Robredo (2003) ao relacionar temas que, em sua opinião, deveriam ser objetos de estudos da Ciência da Informação nos próximos anos, incluiu o estudo do comportamento de diversos tipos e níveis de usuários, nos processos de busca e recuperação da informação, a partir de um dado sistema, como, por exemplo, a Internet e as bases de dados, o que demonstra a pertinência do presente estudo.

O interesse pelos processos de busca e recuperação da informação surgiu após o estágio realizado em um escritório de advocacia na cidade de Marília no período de Julho de

2001 à Julho de 2003, onde, fazendo parte da equipe, na função de estagiário de Biblioteconomia, pode-se realizar, como parte das atividades de intermediário da informação, serviços de buscas de informações em diversas bases de dados disponíveis on-line e em CD-ROM, além do acervo bibliográfico do próprio escritório, recuperando, assim, Leis, Doutrinas e, principalmente, Jurisprudências, ou seja, toda e qualquer informação que trouxesse os subsídios necessários para os advogados solicitantes (sendo que estes se encontram nas cidades de Marília, São Paulo e Rio de Janeiro) defenderem os direitos de seus clientes perante a Justiça. Também, pode-se participar ativamente da idealização, confecção e operacionalização de serviços de informação (do qual as atividades de busca e recuperação da informação são elementos-chave) como o Boletim Jurídico (periódico mensal e destinado aos clientes do escritório) e o e-MURAL (folhetim on-line, de circulação interna), utilizando-se para isso os meios eletrônicos (recursos da Internet), para o desenvolvimento e distribuição de ambos, economizando, assim, enorme tempo para a obtenção de informações relevantes. Todos os serviços prestados durante o estágio tiveram resultados muito positivos em relação a clientes e advogados do escritório.

A partir deste estágio, pode-se perceber a importância e a imprescindibilidade dos serviços de busca de informações, não só para a área do Direito, mas para qualquer área do conhecimento.

Um outro fator que motivou o interesse em investigar este tema, surgiu através do estágio curricular realizado na biblioteca da Unesp-Marília, do qual houve a constatação do acentuado declínio nas solicitações de levantamentos bibliográficos, em bases de dados, junto ao serviço de referência. Conforme o depoimento da bibliotecária responsável por este setor, os alunos da Pós-Graduação eram os principais usuários deste serviço, sendo que no ano de 2003, o número de solicitações diminuiu consideravelmente.



Nos últimos anos a rede UNESP de bibliotecas tem disponibilizado uma série de bases de dados bibliográficas especializadas, que podem ser consultadas em qualquer computador do campus com conexão com a internet. Isto começou a ocorrer, em várias instituições públicas e privadas, a partir do momento em que muitas empresas passaram a oferecer a possibilidade de busca em fontes de informação em versões eletrônicas com interfaces gráficas mais amigáveis (que até então eram exclusivamente impressas, ou compostas por complicadíssimos comandos de busca) para o ambiente acadêmico, iniciando-se uma preocupação das unidades de informação em capacitar o usuário final na realização de suas próprias buscas (MERCADO, 1999).

Assim, pressupõe-se que os usuários têm realizado os seus próprios levantamentos. No entanto, é pertinente verificar como tem se dado este processo de busca a fim de que se possa contribuir para um melhor aproveitamento dessas bases de dados, que é o propósito desta pesquisa. Além disso, um estudo sobre este tema contribuirá para o aperfeiçoamento do serviço de referência e, conseqüentemente, a qualidade e o desenvolvimento das pesquisas realizadas pela Pós-Graduação da UNESP-Marília o que, de modo indubitável, é importante para a própria instituição e para qualquer agência de fomento à pesquisa.

A Faculdade de Filosofia e Ciências (FFC) da UNESP, em Marília, oferece toda a infra-estrutura necessária para a viabilidade desta pesquisa, uma vez que fornece os sujeitos (alunos da Pós-Graduação) para o estudo, possui laboratórios de informática com acesso à Internet; disponibiliza diversas bases de dados on-line nacionais e internacionais, bibliográficas e referenciais, além de todo o apoio dos bibliotecários do serviço de referência da unidade.

Os objetivos desta pesquisa são:

- Identificar o perfil e as necessidades informacionais dos alunos da Pós-Graduação da UNESP – Marília;
- Verificar quais as bases de dados disponíveis para consulta são conhecidas e utilizadas pelos alunos da Pós-Graduação da UNESP – Marília;
- Identificar as necessidades para a otimização dos processos de planejamento e operacionalização das estratégias de busca para uma efetiva recuperação da informação pelos clientes/usuários.

Com os resultados desta pesquisa pretende-se, oferecer subsídios à comunidade bibliotecária para que esta, utilizando-se das tecnologias de informação e comunicação atualmente empregadas, otimize e dinamize seus serviços de busca e referência, habilitando seus clientes/usuários para uma melhor efetividade da recuperação da informação, principalmente a especializada.

A pesquisa que ora se apresenta, inclui o referencial teórico da pesquisa, juntamente com a sua parte empírica organizado, assim, nos capítulos que se segue: **Capítulo 1**, a “**Introdução**”; **Capítulo 2**, intitulado “**Recuperação da Informação**”, que traz uma breve discussão sobre processo de recuperação da informação; **Capítulo 3** “**Recuperação da Informação em Bases de Dados**”, que abrange o tema e suas implicações aos processos de interação entre homem e máquina; **Capítulo 4** “**Serviço de Referência e a Recuperação da Informação pelo Usuário Final**”, que aborda o papel do serviço de referência frente às bases de dados e as interações entre bibliotecário, usuário e bases de dados; **Capítulo 5** “**Procedimentos Metodológicos**” adotados na pesquisa; **Capítulo 6** “**Apresentação e Análise dos resultados**”, no qual os resultados obtidos a respeito do comportamento de busca

dos alunos da Pós-Graduação da UNESP de Marília perante as bases de dados, são descritos e analisados, e o **Capítulo 7** que apresenta as “**Considerações finais**” a respeito da pesquisa e a sugestões e necessidades de outros trabalhos sobre o tema .

## CAPÍTULO 2

Acesso é a palavra chave.  
FIGUEIREDO<sup>2</sup>

### RECUPERAÇÃO DA INFORMAÇÃO

Os desenvolvimentos ocorridos no pós Segunda Guerra Mundial, principalmente os tecnológicos, acabaram por causar novos problemas e necessidades.

Um rápido retrospecto sobre a área da documentação [e da Ciência da Informação] mostra que, nas décadas de 1950 e 1960, com o crescimento do conhecimento científico e tecnológico, houve dificuldades para armazenar e recuperar informações (CINTRA et al, 2002, p.33).

Isto é, o problema crítico do fenômeno conhecido como a explosão da informação, ou seja, “o irreprimível crescimento exponencial da informação e de seus registros, particularmente em ciência e tecnologia” (SARACEVIC, 1996a, p. 42) em meados do século XX, fez, então, com que cientistas, engenheiros, e empreendedores, iniciassem a pesquisar e a trabalhar soluções para esse problema.

Considerando a questão da explosão informacional a Recuperação da Informação veio a ser uma solução bem sucedida encontrada pela Ciência da Informação e está em contínuo processo de desenvolvimento até os dias de hoje (SARACEVIC, 1996a).

Mas, como toda solução faz aparecer seus próprios e específicos problemas, Mooers (1951), questiona:

---

<sup>2</sup> FIGUEIREDO, Nice Menezes de. **Paradigmas modernos da Ciência da informação**. São Paulo: Polis, 1999, 168 p.

- a) como descrever intelectualmente a informação?
- b) Como especificar intelectualmente a busca?
- c) Que sistemas, técnicas ou máquinas devem ser empregados?

E, embora outros problemas tenham surgido, ao longo destes últimos 50 anos, estes três fatores, apontados por Mooers (1951), continuam sendo os principais ou os mais desafiadores até hoje. Saracevic (1996a, p. 44) acrescenta que,

De tais questões, surgiu uma grande variedade de conceitos e construtos teóricos, empíricos e pragmáticos, bem como numerosas realizações práticas. Muitos exemplos históricos podem ilustrar a marcante evolução de sistemas, técnicas e/ou máquinas utilizados para recuperação da informação. Sua variedade vai dos cartões perfurados aos CD-ROMs e acesso on line; dos sistemas não-interativos àqueles de múltiplas possibilidades de interação, com interfaces inteligentes, transformando a recuperação de informação em um processo altamente interativo; de bases documentais para bases de conhecimento; dos textos escritos aos multimídia; da recuperação de citações à recuperação de textos completos; e ainda aos sistemas inteligentes e de respostas a perguntas.

Desde que o termo Recuperação da Informação foi difundido por Calvin Mooers em 1951 tem-se gerado muita polêmica. Hoje, entretanto, ele é amplamente aceito e utilizado pela comunidade científica, ainda que existam diversos pontos de vista do que seja “Recuperação da Informação”.

Como define Robins (2000, p. 57, tradução nossa),

A recuperação da informação é uma disciplina interessada nos processos pelos quais questões são apresentadas a sistemas de informação [...]. O resultado final desses processos é uma lista de documentos que são um subconjunto dos sistemas de informação. O processo pode ser realizado por qualquer meio, mas essencialmente, quando atributos específicos de uma questão são correspondidos com atributos específicos de um documento, o documento é incluído na lista<sup>3</sup>.

Para que o usuário encontre mais facilmente a informação de que necessita, acredita-se que um sistema de Recuperação da Informação deveria ser capaz de realizar o armazenamento, a recuperação e a manutenção da informação (KOWALSKI, 1997). Ou deveria tratar de catalogação, indexação, classificação e busca de grande quantidade de informações, particularmente na forma textual (STRZALKOWSKI, 1999). Ou ainda, além de armazenar, deveria representar, organizar e dar acesso a itens de informação (textos, documentos, etc) (BAEZA-YATES, 1999).

No entanto, autores como Frantz, Shapiro, Voiskunskii (1997) salientam que um sistema de Recuperação da Informação não recupera informação, pois a informação consiste no relacionamento que ocorre entre o usuário e os símbolos que recebe. Assim, pelo caráter extrínseco da informação focado pela teoria dos sistemas (MACIEL, 1974), quem a pode recuperar é o usuário e não o sistema. Como analisa Braga (1995, p. 85),

---

<sup>3</sup> Information retrieval (IR) is a discipline concerned with the processes by which queries presented to information systems [...]. The end result of the matching process is a listing of texts that are a subset of the total store. Any number of means may accomplish the matching process, but essentially, when specified attributes in a query are found to correspond with specified attributes of a text, the text is included in the listing.

Na verdade, os Sistemas de Recuperação da Informação não recuperam informação, ou recuperam apenas uma informação potencial, uma probabilidade de informação, que só vai se consubstanciar a partir do estímulo externo documento, se também houver uma identificação (em vários níveis) da linguagem desse, documento, e uma alteração, uma reordenação mental do receptor-usuário.

Assim, embora o acesso à informação tenha sido potencializado pelo desenvolvimento das tecnologias de informação e comunicação, há uma necessidade mais acentuada de se dominar o processo de busca e recuperação da informação, que inclui as atividades de escolha dos termos que melhor representem os conceitos que estão sendo buscados; a elaboração de uma estratégia de busca eficaz e o conhecimento dos recursos oferecidos pelos sistemas de busca (ROWLEY, 2002), para se extrair a informação mais desejável.

De certa forma a Recuperação da Informação é atualmente a atividade principal (ou ao menos deveria ser), em pesquisas, na área da Biblioteconomia e da Ciência da Informação, pois “a origem e antecedentes sociais da ciência da informação evoluíram para a recuperação da informação” (SARACEVIC, 1995 citado por ROBREDO, 2003, p. 65), propondo uma solução com a mudança do enfoque e da conceituação da recuperação da informação. “Com efeito, foi abandonada a perspectiva preferencial de recuperação bibliográfica e normalização classificatória e descritiva, buscando-se a construção de linguagens próprias” (CINTRA et al, 2002, p. 33) como, por exemplo, o desenvolvimento da indexação ou seja,

[...] o trabalho com a recuperação da informação foi também o responsável pelo desenvolvimento da Ciência da Informação como um campo onde se

interpenetram os componentes científicos e profissionais. Certamente, a recuperação da informação não foi a única responsável pelo desenvolvimento da Ciência da Informação, mas pode ser considerada como principal; ao longo do tempo a Ciência da Informação ultrapassou a recuperação da informação, mas os problemas principais tiveram sua origem aí e ainda constituem seu núcleo (SARACEVIC, 1996a, p.45).

Isto é, um dos maiores problemas da Biblioteconomia e da Ciência da Informação é a Recuperação da Informação, visto que o aumento exponencial do uso, desenvolvimento e aplicação das tecnologias de informação e comunicação possibilitou a criação de diversos novos serviços, produtos, sistemas e redes de informação de forma que uma crescente massa documental e informacional vem sendo constantemente disponibilizada na Internet por meio de diversas bases e bancos de dados on line. Saracevic (1996a, p. 45) ainda acrescenta que,

Bastante significativa, também, foi a emergência do pragmatismo na aplicação empresarial da recuperação da informação: a indústria informacional ou, para ser mais preciso, o setor que lida com a criação e distribuição de bases de dados e de serviços on line decorrentes, bem como o acesso à informação e sua disseminação. Essa indústria da informação tem suas raízes diretamente relacionadas com os trabalhos de recuperação da informação dos anos 50 e 60, que culminaram com a emergência de serviços on line nos 70 e com a viabilização internacional da indústria da informação nos anos 80.

Ou seja, a recuperação da informação também influenciou a forma e a evolução da indústria informacional.



Diante dessas idéias iniciais, vemos que, a Recuperação da Informação engloba os aspectos intelectuais da descrição de informações e suas especificidades para a busca, além de quaisquer sistemas, técnicas ou máquinas empregados para o desempenho da operação (MOOERS, 1951).

## **2.1 Mas o que seria Informação ?**

A definição de um conceito de informação é bastante complexa, pois ele vem sendo empregado nas situações mais diversas, como aponta Robredo (2003, p. 1),

Embora o termo informação seja, nos tempos que correm, utilizado com avassaladora freqüência, se constituindo numa presença constante, quase obsessiva, tanto no vocabulário técnico-científico quanto na fala de uma grande parte das pessoas comuns, integrantes de todas as classes sociais, nada é menos seguro que esse vocábulo possua o mesmo significado para todos os que o utilizam ou dele ouvem falar.

A chamada era da informação, ou a sociedade da informação, ou ainda a explosão informacional ocorrida em meados do último século, carregam em si a idéia de que tudo e todos contêm e processam informação.

Autores como Ribeiro e Santos (2003, p. 128) dizem que Informação é um “conceito genérico de tudo que possa representar notícia, conhecimento ou comunicação”. Já para Shapiro e Varian (1999, p.15), “em essência qualquer coisa que puder ser digitalizada – codificada como um fluxo de bits - é informação”.

Segundo Cohen (2002, p.27),

Para a análise informacional, a informação é um agente dissipador de incertezas cujo objetivo é proporcionar alterações no comportamento das pessoas, reduzindo a incerteza. Diferencia-se da informática e da cibernética. A informática tem o objetivo de tratar da automação da informação, e a cibernética volta-se para os problemas de controle da informação nas máquinas e nos organismos vivos.

Sant’Ana e Santos (2004, p. 55) consideram informação

[...] como um conjunto finito de dados dotado de semântica e que tem a sua significação ligada ao contexto do agente que a interpreta ou recolhe e de fatores como tempo, forma de transmissão e suporte utilizado. O valor desse conjunto poderá diferir da soma dos valores dos dados que o compõem, dependendo do processo de contextualização no agente que o recebe.

Muitos pesquisadores e estudiosos discutem o que é informação, cada qual atribuindo sentidos e significados diferentes. No entanto este estudo restringiu-se aos conceitos sobre informação no campo de atuação da Biblioteconomia e da Ciência da Informação onde a informação pode ser “um conhecimento inscrito (gravado) sob a forma escrita (impressa ou numérica), oral ou audiovisual” (LE COADIC, 1996, p.5).

Le Coadic (1996, p.5) ainda acrescenta que,

A informação comporta um elemento de sentido. É um significado transmitido a um ser consciente por meio de uma mensagem inscrita em um suporte espacial-temporal: impresso, sinal elétrico, onda sonora, etc. Essa inscrição é feita graças a um sistema de signos (a linguagem), signo este que é um elemento da linguagem que associa um significante a um

significado: signo alfabético, palavra, sinal de pontuação. [...] o objetivo da informação permanece sendo a apreensão de sentidos ou seres em sua significação, ou seja, continua sendo o conhecimento; e o meio é a transmissão do suporte, da estrutura.

Mas, por outro lado, a combinação de um estímulo externo, uma reordenação mental (classificação) e uma designação (ainda que articulada apenas em nível de identificação de algo que não o havia sido anteriormente) podem ser vista como uma primeira aproximação ao conceito de informação. Esta noção é, de certa forma, aquela que vê a informação como aquilo que é capaz de transformar estruturas. Essa visão é, no entanto, ainda muito ampla e abrange várias instâncias em que o termo informação não é empregado. A idéia central do conceito apresentado é a de transformação de estruturas (BRAGA, 1995). Ou seja, a informação é uma representação simbólica e formal de fatos ou idéias, potencialmente capaz de alterar o estado de conhecimento (GONÇALVES, 1996).

COHEN (2002, p. 27) acrescenta ainda que,

A contextualização é a principal característica para a informação. Uma informação pode ser importante, ter valor, fazer sentido para uma pessoa, e não causar nenhuma mudança em outra pessoa. A informação, por fim, representa algo que existe. A informação, numérica ou não, não é uma abstração, mas está inexoravelmente ligada a uma representação física [...]. Assim, o processamento da informação está inexoravelmente ligado ao universo material e suas leis.

Segundo Robredo (2003, p. 9), a informação, no âmbito da Biblioteconomia e da Ciência da Informação, é suscetível de ser

- Registrada (codificada) de diversas formas,
- Duplicada e reproduzida *ad infinitum*,
- Transmitida por diversos meios,
- Conservada e armazenada em suportes diversos,
- Medida e quantificada,
- Adicionada a outras informações,
- Organizada, processada e reorganizada segundo diversos critérios,
- Recuperada quando necessário segundo regras preestabelecidas.

Neste sentido, inclui-se o registro e transmissão do conhecimento, o armazenamento, processamento, análise organização e recuperação da informação registrada e os processos e técnicas relacionados, não se esquecendo que,

Outras noções possíveis de informação [...] enfocam a informação como redutora de incertezas, entropia negativa, fator de homeostase, força básica, utilidade pública, algo que é transmitido em um processo de comunicação, uma “coisa” (BRAGA, 1995, p.2).

## **2.2 Por que Recuperar a Informação ?**

No presente mundo globalizado, torna-se imprescindível uma constante atualização, objetivando sempre, as melhores condições para garantir a sobrevivência em um mercado altamente competitivo, sendo que “a obtenção ágil da informação e, principalmente, da informação qualificada, é o grande desafio dos nossos tempos” (JAMIL; NEVES, 2000).

De acordo com Castells (1999, p. 87),

Uma nova economia surgiu em escala global nas duas últimas décadas. Chamo-a de informacional e global para identificar suas características fundamentais e diferenciadas e enfatizar sua interligação. É informacional porque a produtividade e a competitividade de unidades ou agentes nessa economia (sejam empresas, regiões ou nações) dependem basicamente de sua capacidade de gerar, processar e aplicar de forma eficiente a informação baseada em conhecimentos. É global porque as principais atividades produtivas, o consumo e a circulação, assim como seus componentes (capital, trabalho, matéria-prima, administração, informação, tecnologia e mercados) estão organizados em escala global, diretamente ou mediante uma rede de conexões entre agentes econômicos. É informacional e global porque, sob novas condições históricas, a produção é gerada, e a concorrência é feita em uma rede global de interação. E ela surgiu no último quartel do século XX por que a Revolução da Tecnologia da Informação fornece a base material indispensável pra essa nova economia. É a conexão histórica entre a base informação/conhecimentos da economia, seu alcance global e a revolução da Tecnologia da Informação que cria um novo sistema econômico distinto [...].

Assim, um dos mais importantes recursos das organizações, nos dias atuais, é a Informação que, deixou de ser componente para ser ferramenta de ação (JAMIL; NEVES, 2000).

Visto que a Ciência é tributária do tempo e do espaço (MEADOWS, 1999), estar atualizado é essencial não apenas para o desenvolvimento de pesquisas, mas para qualquer meio ou instituição. No ambiente acadêmico, por exemplo, a informação que alimenta a produção científica é a informação resultante de outras pesquisas, que em geral são

divulgadas através de periódicos, teses, dissertações, livros, entre outros. Mesmo se tratando de um meio bastante seletivo e rigoroso, a quantidade de informações disponíveis esta além da capacidade de absorção humana, ainda que se restrinja a uma área bastante específica. É neste sentido que uma recuperação da informação com eficácia torna-se importante.

Desta forma, para informarem-se sobre os acontecimentos políticos, econômicos, sociais ou culturais; ou sobre as mais novas inovações tecnológicas; ou ainda, sobre os acontecimentos mais recentes no mundo dos negócios de maneira eficaz, é necessário que empresas, instituições, Universidades, ou qualquer outro tipo de organização, assimilem os mecanismos (ou seja, sistemas de informação, pessoal qualificado, etc) que ofereçam os recursos e subsídios necessários para as tomadas de decisões e, assim, realizarem suas manobras, no tempo mais hábil possível.

### **2.3 A Recuperação da Informação e as Tecnologias da Informação e Comunicação**

Economizar tempo nos dias de hoje, mais do que nunca tem adquirido enorme valor, pois vivemos no liminar da era da informação e a própria sobrevivência humana tem exigido uma profunda atualização informacional. Em todas as atividades profissionais, a velocidade no acúmulo de conhecimento está aumentando aceleradamente. Surgiu uma sociedade baseada na tecnologia da informação ou, como aponta Castells (1999, p. 49), “[...] estamos vivendo um desses raros intervalos na história. Um intervalo cuja característica é a transformação de nossa ‘cultural material’ pelos mecanismos de um novo paradigma tecnológico que se organiza em torno da tecnologia da informação”. Isto é, todas as atividades relacionadas com a manipulação, a edição, o armazenamento, a distribuição e recuperação da informação, assim como todas as formas de trabalho que lidem diretamente com dados textuais, simbólicos, numéricos, visuais e até mesmo auditivos, precisam agora adequar-se à forma digital (LEVACOV, 1997).

Com o advento das novas tecnologias de informação e comunicação, a Internet eliminou, consideravelmente, o incomodo de ter de aguardar as correspondências trazidas pelos correios, por exemplo. Segundo Cohen (2002, p. 26),

[as] novas tecnologias a cada dia possibilitam melhorias que diminuem obstáculos como tempo e espaço. Tanto as pessoas quanto as organizações têm trocado mais informações, seja pelo envio de mensagens ou troca de arquivos.

Stix (2002, p. 51) diz ainda que,

hoje, a quantidade de eventos comprimidos em determinado intervalo de tempo – seja um ano ou um nanossegundo – aumenta sem parar. A era tecnológica [e da informação] tornou-se um jogo em que todos sempre querem [e precisam] estar à frente.

Ou seja, as tecnologias de informação e comunicação cada vez mais vêm proporcionando e estabelecendo um ritmo universalmente mais acelerado para todos, no qual se torna fundamental a constante atualização, seja qual for o tipo de organização.

A difusão acelerada das novas tecnologias de informação e comunicação vem promovendo profundas transformações na economia mundial e está na origem de um novo padrão de competição globalizado, em que a capacidade de gerar inovações em intervalos de tempo cada vez mais reduzidos é de vital importância para empresas e países (TAKAHASHI, 2000, p.17).

Os computadores e a infra-estrutura para o acesso a informação remota tornaram-se mais rápidos e eficientes, inclusive em sua capacidade de armazenamento da informação. Assim, é possível acessar diversas bases de dados de qualquer lugar do mundo que esteja conectado a Internet. A interface gráfica da Web, por sua vez, tornou o processo de busca mais amigável, principalmente, “[...] as interfaces de busca, dirigidas ao usuário final”<sup>4</sup> (ANGÓS ULLATE; SALVADOR OLIVÁN; FERNÁNDEZ RUÍZ, 2001, p. 141, tradução nossa).

Além disso, um dos grandes benefícios, com desenvolvimento das tecnologias de informação e comunicação, é a redução dos custos de produção e distribuição dos documentos. Segundo Levacov (1997), muitas são as vantagens, entre elas, o custo reduzido, a velocidade do processo, a amplitude potencial de acesso e disseminação, a inclusão de dados baseados em tempo (vídeo, áudio, animações), a facilidade de atualização e/ou inclusão de novos dados as funções de busca e indexação dos artigos e a rede hipertextual com âncoras para outros documentos afins.

Robredo (2003, p.23), nesta mesma linha de pensamento ainda destaca,

O barateamento dos custos da codificação e o aumento no nível de segurança, trazidos pelos avanços das tecnologias da informação e da comunicação, muito estão contribuindo para a rápida e cada dia mais abrangente mudança de práticas e hábitos relacionados com a armazenagem, organização e conservação de grandes volumes de registros do conhecimento (digitalização, grandes bases de dados), com a transmissão, transferência e acesso às fontes originais (bibliotecas virtuais, grandes enciclopédias em linha), socialização da cultura, do conhecimento e

---

<sup>4</sup> [...] los interfaces de búsqueda, dirigidos al usuario final.



da educação (museus virtuais, ensino a distância), assim como do acesso a informações políticas, legais, bancos em linha, turismos) e muito mais.

Por conseguinte, vê-se que atualmente há uma crescente massa documental e informacional sendo disponibilizada na Web por meio de diversas bases e bancos de dados, conseqüência, também, da crescente preocupação de automação de vários segmentos como empresas, órgãos públicos e privados, instituições científicas e, principalmente, Universidades e órgãos governamentais (sejam eles federais, estaduais ou municipais), por exemplo, onde

O *governo*, nos níveis federal, estadual e municipal, tem o papel de assegurar o acesso universal às tecnologias de informação e comunicação e a seus benefícios, independentemente da localização geográfica e da situação social do cidadão, garantindo níveis básicos de serviços, estimulando a interoperabilidade de tecnologias e de redes (TAKAHASHI, 2000, p. 11, grifo do autor).

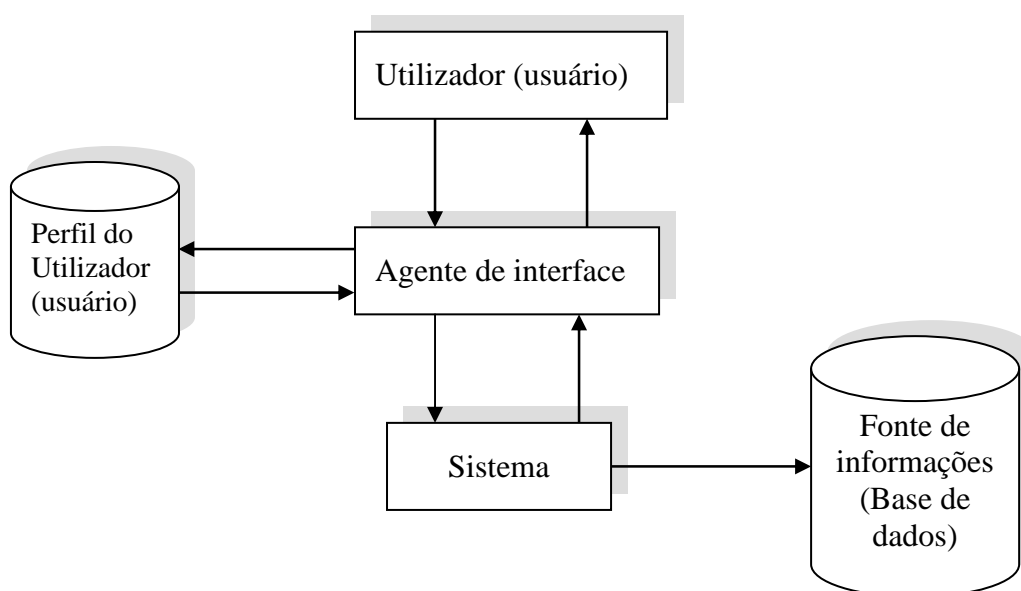
Com a automação, o intuito das instituições, públicas ou privadas, é melhorar as suas atuações e serviços junto à sociedade em geral ou a qual ela se destina, na mesma medida em que as tecnologias de informação e comunicação contribuem para oferecer maior transparência à administração das diversas instituições. Logo vê-se que,

Assim, a Internet vai evoluindo como repositório mundial de informações. Conseqüentemente, gama variada de sistemas e serviços emerge, disponibilizando massa informacional de forma mais acessível. Alcança espaços maiores e diferentemente abertos, num sentido globalizante, principalmente por inovar a extensividade; privilegia também o usuário

remoto, o conhecimento de novas culturas e soluções atualizadas a muitos profissionais para a melhoria das atividades (MODESTO; MACEDO, 1999, p. 63).

Os avanços e desenvolvimentos das tecnologias de informação e comunicação são amplos. Exemplo disto, é o que já vem ocorrendo com os sistemas comerciais de venda pela internet e outros, onde o perfil do cliente é constantemente analisado, durante toda a interação cliente-sistema, capturando suas preferências por produtos e/ou serviços, “desenhando” o perfil deste cliente para ofertá-lo com anúncios de produtos e/ou serviços personalizados, através da interface Web (VIDOTTI, 2004, informação verbal<sup>5</sup>).

Do mesmo modo o perfil do cliente/usuário de uma unidade de informação é organizado de forma a armazenar suas preferências formais, identificação e o conhecimento sobre as interações com o sistema de recuperação (RIBEIRO JÚNIOR, 2003b, p. 21).



**Figura 1:** Arquitetura genérica dos agentes de filtragem de informação  
**Fonte:** NOGUEIRA, 2000, p. 14

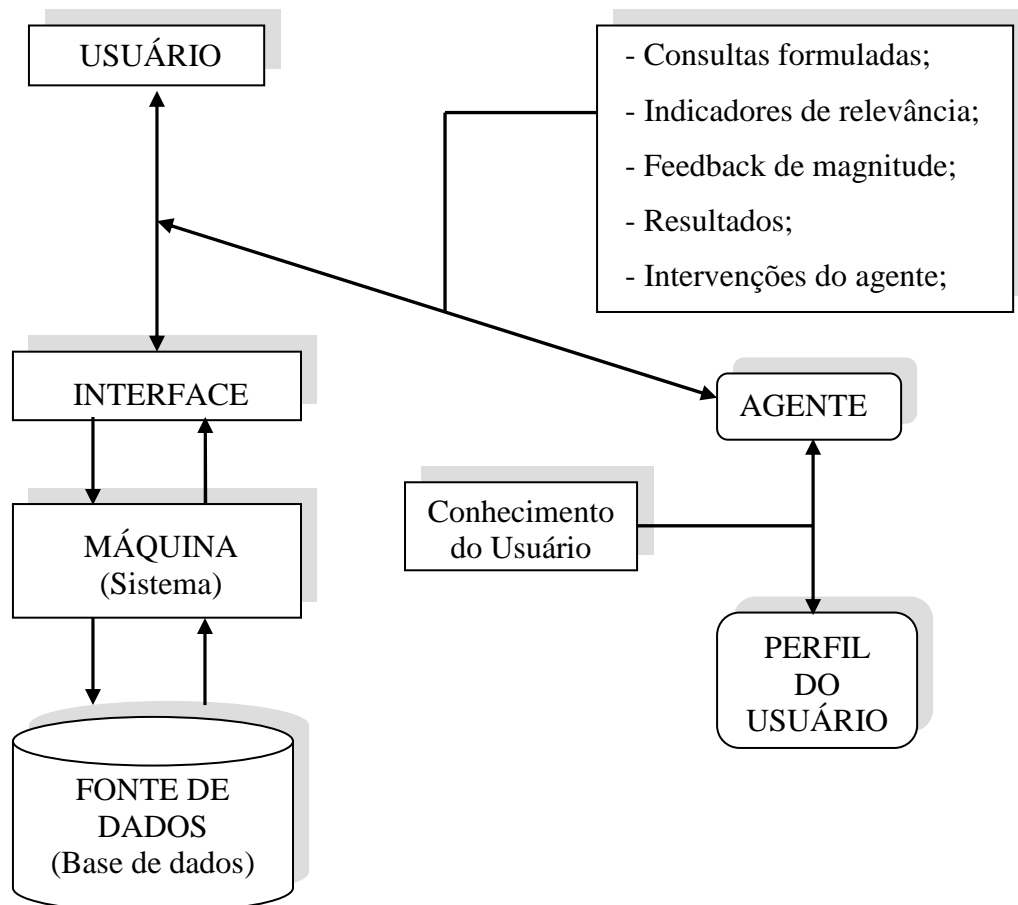
<sup>5</sup> VIDOTTI, Silvana Aparecida Borsetti Gregório. **Win Isis gerenciador de bases de dados**. Marília, Unesp, 2004. Notas de aula.

Segundo Nogueira (2000) o sistema interage com o usuário/cliente através da interface que o agente<sup>6</sup> de interface disponibiliza. Assim, o agente de interface é, logicamente, o componente do sistema que tentará construir um perfil do usuário/cliente. Este perfil deve ser baseado no comportamento do usuário/cliente nos termos da interação com a interface, e do conhecimento do agente em relação a semântica ligada aos componentes individuais da interface.

Tal tecnologia é possível devido ao desenvolvimento dos **sistemas ou agentes inteligentes**. É possível encontrar na literatura diversos trabalhos que envolvem a recuperação de informação e “sistemas inteligentes”, incluindo a abordagem de “agentes inteligentes”. A aplicação destes sistemas, é baseada na expansão da consulta através da associação de termos relacionados aos mesmos conceitos de interesse do usuário, buscando aumentar a performance em consultas realizadas.

---

<sup>6</sup> Um programa de computador ou processo que opera sobre uma aplicação cliente ou servidor e realiza uma função específica, como uma troca de informações, através da manipulação de dados (RIBEIRO; SANTOS, 2003).



**Figura 2:** Ambiente e seus principais componentes  
**Fonte:** RIBEIRO JÚNIOR, 2003b

Segundo Haverkamp e Gauch (1998), a expansão da consulta é feita por meio do auxílio de um agente, que irá interagir em um ambiente onde se encontram os participantes do processo de recuperação de informação.

Neste ambiente (figura 2) existe o usuário, que interage com o sistema de recuperação (através de uma **interface**), o agente, que interage com o usuário e o sistema de recuperação, atuando como mediador e também interagindo com o **conhecimento** acumulado sobre os processos de recuperação de instâncias anteriores deste ambiente, armazenado na forma de um **perfil**. O agente atua (**intervem**) sobre o fluxo de comunicação existente entre o sistema de recuperação e o usuário. Ele obtém das **consultas formuladas** e dos documentos considerados **relevantes** pelo usuário, o **feedback de magnitude**, trazendo os **resultados** (documentos) e, quando existirem, os descritores dos documentos recuperados. No seu escopo

de atuação o agente deverá buscar, no perfil do usuário, descritores de outros documentos cuja representação semântica se assemelha com a dos utilizados pelo usuário naquela instância de recuperação. Como o seu discurso pode se modificar com o tempo, o que se verifica normalmente à medida que o usuário aprende mais sobre o sistema e o domínio de conhecimento em questão, o agente continua aumentando as associações semânticas entre o sistema de recuperação e a linguagem do usuário (HAVERKAMP; GAUCH, 1998).

Assim, Instituições não-comerciais, como bibliotecas ou o governo (e todos os seus serviços e organizações), podem usá-los para estender os seus serviços a um maior grupo de usuários (na Internet ou nas suas redes locais) (NOGUEIRA, 2000).

Desta forma, a Recuperação da Informação esta intimamente ligada às tecnologias de informação e comunicação uma vez que “a tecnologia da informação centrada no computador, redes e periféricos não serviu, apenas, para acelerar os velhos processos, mas para muda-los” (MILANESI, 2002, p. 08).

Mas, como diz Pitassi e Leitão (2002, p.78),

O Mito criado em torno dos benefícios que a tecnologia em geral e a Tecnologia da Informação (TI) em particular inexoravelmente trazem para a humanidade deixa nos mais incautos a sensação de estarmos no limiar de uma nova sociedade, já denominada “economia da informação”, inerente ao que se esta se chamando de “sociedade do conhecimento”. Nessa perspectiva, assume-se o pressuposto de que a tecnologia resolverá os problemas da humanidade. No entanto, a tecnologia não é capaz de determinar nada por si só, pois é utilizada dentro de um contexto político-ideológico mais amplo.

Observa-se aqui, portanto, que as mudanças atribuídas às tecnologias de informação e comunicação, fundamentam-se basicamente na velocidade e na digitalização, na conectividade e no “acesso universal” à informação e,

[...] servirão como instrumentos para aumentar a competitividade das organizações, jamais para mudar os paradigmas da sociedade em que se inserem. São, portanto, mudanças adaptativas, que visam à manutenção do status quo, ainda que venham embaladas por um discurso pseudo-revolucionário e repleto de clichês cibernéticos (PITASSI; LEITÃO, 2002, p. 80).

Vê-se então que, as interações dos usuários com as tecnologias de recuperação da informação, dentro de um contexto mais holístico do comportamento da informação humana, devem fazer parte do atual conceito da Biblioteconomia e da Ciência da Informação (SPINK, 2000).

## **2.4 As Bases de dados**

A Recuperação da Informação tem sido considerada como uma expressão que engloba desde a recuperação de documentos (ou de textos) até a recuperação de conhecimento (LEWIS; JONES, 1996). Entretanto, tem sido mais freqüentemente tratado como sinônimo de “recuperação de documentos” e de “recuperação de textos” (SPARCK-JONES; WILLET, 1997), acreditando-se que sua tarefa essencial seja recuperar documentos ou textos com informação. Ou seja, ao entender que o foco de um sistema de Recuperação da Informação é a informação textual, sendo que, “o termo textual pode ser substituído ainda por imagens, sons,

vídeos, pinturas, ou qualquer outro artefato de atividade intelectual”<sup>7</sup> (ROBINS, 2000, p.57, tradução nossa), vê-se que,

O processo de recuperação da informação consiste em identificar, no conjunto de documentos (corpus) de um sistema, quais atendem à necessidade de informação do usuário. O usuário de um sistema de recuperação de informação está, portanto, interessado em recuperar “informação” sobre um determinado assunto e não em recuperar dados que satisfazem sua expressão e busca, nem tampouco documentos, embora seja nestes que a informação estará registrada. Essa característica é o que diferencia os sistemas de recuperação de informação dos Sistemas Gerenciadores de Bancos de Dados (ou simplesmente “banco de dados”), estudados e implementados desde o nascimento da Ciência da Computação (FERNEDA, 2003, p. 14).

Para Sant’na e Santos (2004, p. 55), tratam “[...] o termo dado como um elemento básico, formado por signo ou conjunto finito de signos que não contém, intrinsecamente, um componente semântico, mas somente elementos sintáticos”.

Logo, a Recuperação da Informação não é a recuperação de dados, sendo válido supor que um Sistema de Recuperação da Informação seja diferente de um sistema de gerenciamento de banco de dados (SPARCK-JONES; WILLET, 1997) onde pode ser armazenados conjuntos de “informações” (dados) de forma computadorizada, através de sistemas de números binários (por exemplo), representados por pequenos sinais eletrônicos –

---

<sup>7</sup> [...] the term text may be substituted with still images, sounds, video clips, paintings, or any other artifact of intellectual activity.

“ativado” ou “desativado” (um ou zero) – apenas legíveis por máquinas (RIBEIRO; SANTOS, 2003).

Segundo Tramullas Saz (1997, p. 73, tradução nossa),

A perspectiva tradicional das Ciências da Informação e da Documentação, tem favorecido a divisão entre “recuperação de dados” e “recuperação da informação” [...]. A primeira delas se centrou em ambientes em que a informação se encontrava altamente estruturada, como nas bases de dados relacionais, enquanto que a segunda lidou com ambientes muito pouco estruturados, como por exemplo, grandes quantidades de informação textual. Teoricamente, as tarefas de recuperação, na primeira delas, limitaram sua dificuldade ao correto uso de sofisticadas, e altamente formalizadas, linguagens de recuperação, enquanto que na segunda influíram numerosos fatores intrínsecos e extrínsecos, como a linguagem natural, os mecanismos de indexação, etc.<sup>8</sup>

Os sistemas gerenciadores de bancos de dados armazenam a informação de forma estruturada. Já os sistemas de recuperação de informação, armazenam a informação não estruturada, ou seja, aquela informação contida nos documentos. Neste sentido Ferneda (2003, p. 15) destaca que,

---

<sup>8</sup> La perspectiva tradicional de las Ciencias de la Información y la Documentación, ha favorecido la división entre “data retrieval” e “information retrieval”, [...]. La primera de ellas se centraba en entornos en los que la información se encontraba altamente estructurada, como en las bases de datos relacionales, mientras que la segunda lidiaba con entornos muy poco estructurados, como por ejemplo, grandes cantidades de información textual. Teóricamente, las tareas de recuperación, en la primera de ellas, limitaban su dificultad al correcto uso de perfeccionados, y altamente formalizados, lenguajes de recuperación, mientras que en la segunda influían numerosos factores intrínsecos y extrínsecos, como el lenguaje natural, los mecanismos de indización, etc.



Os sistemas de bancos de dados têm por objetivo a recuperação de todos os objetos ou itens que satisfazem precisamente às condições formuladas através de uma expressão de busca. Em um sistema de recuperação de informação essa precisão não é tão restrita. A principal razão para esta diferença está na natureza dos objetos tratados por estes dois tipos de sistema. Os sistemas de recuperação de informação lidam com objetos lingüísticos (textos) e herdam toda a problemática inerente ao tratamento da linguagem natural. Já um sistema de bancos de dados organiza itens de “informação” (dados), que têm uma estrutura e uma semântica bem definidas. Os sistemas de informação podem se aproximar do padrão que caracteriza os bancos de dados na medida em que sejam submetidos a rígidos controles, tais como vocabulário controlado, listas de autoridades, etc.

Difícil falar-se, nos dias de hoje, em recuperação da informação sem citar as bases e bancos de dados eletrônicos, ou seja, grandes quantidades de dados armazenados em meios ópticos ou magnéticos embora, no nosso dia-a-dia, as bases de dados serem muito comuns, raramente, paramos para pensar nisso. As listas telefônicas, os dicionários, nossa agenda telefônica, as enciclopédias entre outros, são todos exemplos de bases de dados.

Na literatura especializada há uma certa confusão a respeito de bases e bancos de dados, onde muitas vezes são tratados como sinônimos. Ribeiro e Santos (2003, p. 23) definem base de dados como um,

Cadastro de dados armazenados em meio magnético [ou óptico] estruturado com seus respectivos dicionários, campos recuperáveis e formatos de saída predefinidos, apoiado em software de um sistema de computador; [...]

informações ou arquivos guardados em um computador para recuperação e uso.

Para Rowley (2002, p. 125), “Uma base de dados é uma coleção geral e integrada de dados junto com a descrição deles, gerenciada de forma a atender a diferentes necessidades de seus usuários”

Já Cianconi (1987, p. 54) afirma que,

[...] segundo a maioria da literatura na área de Informação, pode-se conceituar Base de Dados como um conjunto de dados interrelacionados, organizados de forma a permitir recuperação de informações. [E que] Banco de Dados, embora freqüentemente encontrado como sinônimo de base de dados, pode ser visto como um conjunto de bases de dados.

Heemann (1997, p. 14) em sua dissertação de mestrado segue as conceituações em que tratam as bases de dados como armazenadoras de registros e/ou dados visando sua **recuperação**. Segundo a autora “essas abordagens são as mais utilizadas na Ciência da Informação, e pelos usuários em geral, para se referir a esse tipo de aplicativo”.

Há ainda autores, como Menou e Guinchat (1994, p. 295), que consideram bases e bancos de dados termos e conceitos diferentes. Para eles: “uma base de dados [bibliográfica] é um conjunto organizado de referências bibliográficas de documentos que se encontram armazenados fisicamente em vários locais”. Segundo estes autores, “Um usuário que busca uma informação em uma base de dados bibliográfica recupera uma referência [...]”. Já “os bancos de dados tratam das informações factuais, numéricas ou textuais diretamente utilizáveis”. Ou seja, segundo os autores, os bancos de dados, proporcionam acesso direto às informações primárias (textos completos).

Ou ainda, de acordo com Lancaster, (1993, p. 305), pode-se conceituar base de dados bibliográficas como:

Uma coleção de itens sobre os quais podem ser realizadas buscas com a finalidade de revelar aqueles que tratam de um determinado assunto. A base de dados consiste em artefatos, como livros (o acervo de uma biblioteca é uma base de dados com certeza), ou registros que representam os artefatos, como por exemplo, registros bibliográficos constantes de páginas impressas, de fichas ou de meios eletrônicos.

Há diversas outras definições sobre bases e bancos de dados. Entretanto ficaremos com aquelas que são regidas pela Biblioteconomia e Ciência da Informação como as definidas por Cianconi (1987), Heemann (1997) e Rowley (2002).

No âmbito das instituições acadêmicas (o qual é o interesse desta pesquisa), os tipos de bases de dados oferecidas, de acordo com Rowley (2002) são:

- Bases de dados bibliográficos: incluem citações ou referências bibliográficas e, às vezes, resumos de trabalhos publicados em periódicos, livros, jornais, relatórios, patentes, teses ou anais de congresso.
- Bases de dados catalográficos: mostram o acervo de determinada biblioteca ou rede de bibliotecas. Comumente, bases de dados de livros, periódicos e outros itens que uma biblioteca possui em seu acervo.

Existem ainda as **bases de dados textuais** (ou de fontes) que contêm os dados originais e constituem um tipo de documento eletrônico (ROWLEY, 2002), “Contém a

informação completa (ou dados primários), os quais não requerem do usuário ir a outras fontes para obter respostas” (RIBEIRO; SANTOS, 2003, p.27), ou seja, neste tipo de base de dados, recupera-se o texto completo dos documentos.

Esse tipo de base de dados vem se tornando cada vez mais comum, uma vez que

O surgimento das novas tecnologias de informação permitiu a otimização da produção, acesso e disseminação da informação, mudando o conceito tradicional de informação bibliográfica baseada em documentos impressos. O acesso, via Internet, a novos recursos informacionais, como hipertexto, hipermídia, listas de discussão, conferências, virtuais, além da versão eletrônica de documentos impressos, tem se tornando uma realidade cada vez mais presente no dia-a-dia dos profissionais da informação (CORREA DA CRUZ et al, 2003, p.47).

Exemplo disto é o constante desenvolvimento das bibliotecas digitais e das iniciativas dos repositórios de arquivos abertos para a publicação editorial acadêmica.

As bases de dados, como já dito anteriormente, podem ser armazenadas em meios ópticos ou magnéticos e acessadas localmente (em uma Intranet ou em CD's ROM, por exemplo) ou remotamente pela Internet.

Segundo Mercado (1999, p. 260, tradução nossa), durante o final dos anos 1980 e início dos anos 1990 “os Discos Compactos foram usados para armazenar informação em base de dados e serem periodicamente atualizados através de discos de substituição”<sup>9</sup>. E, ao mesmo tempo, “um número crescente de bibliotecas começou a automatizar seus catálogos

---

<sup>9</sup> Compact disks were used to store database information and were used to store database information and were periodically updated by replacement disks.

durante a década de 1980 para ser catálogos de acesso de público on-line (OPACs)”<sup>10</sup> (MERCADO, 1999, p. 261, tradução nossa).

Porém, conforme ressalta Rowley (2002, p. 166), “os sistemas informatizados de recuperação da informação, em comum com muitos outros sistemas, passaram por várias gerações nos últimos 20 anos [...]”. Logo abaixo, apresentam-se as gerações de sistemas de recuperação da informação:

<b>Primeira geração</b>	<b>Metadados</b>	Interfaces baseadas em comandos, usuários especialistas e intermediários; número limitado de sistemas em linha nas instituições e disponíveis externamente por meio de serviços de busca em linha.
<b>Segunda geração</b>	<b>Dados com texto integral</b>	Interfaces baseadas em menus e comandos; recursos de recuperação adicionais, como hipertexto e buscas em texto completo; interfaces baseadas em DOS; previsto o acesso pelo usuário final, mas nem sempre possível ou alcançado; sistemas em linha, com os primeiros sistemas baseados em cederrom.
<b>Terceira geração</b>	<b>Multimídia</b>	Interfaces gráficas; foco no acesso pelo usuário final; orientada para o mercado e com ênfase em pacotes de produtos; armazenamento e distribuição em cederrom ou em redes de alta capacidade; multimídia; intermediário com a função de instrutor; maior uso no lar e em ambientes de acesso público.

**Figura 3:** Gerações de sistemas de recuperação da informação

**Fonte:** Rowley, 2002, p. 167

Foi quando os computadores se tornaram mais rápidos e mais poderosos que,

[...] a Internet desenvolveu-se do texto-básico à interface de gráficos da World Wide Web. Esta informação tornou-se totalmente acessível com o uso de softwares gráficos tal como o Mosaic e o Netscape, os quais

<sup>10</sup> An increasing number of libraries began to automate their catalogs during the 1980s to be online public access catalogs (OPACs).

organizavam a informação em uma atraente e fácil forma de navegação (MERCADO, 1999, p. 261)<sup>11</sup>.

Assim, os sistemas tornaram-se mais acessíveis para o usuário inexperiente e ocasional.

“Até a poucos anos, todos os sistemas de recuperação da informação baseavam-se em comandos. Tais sistemas eram considerados impenetráveis pelo usuário inexperiente” (ROWLEY, 2002, p. 167), além de que cada sistema adotava uma linguagem de comando diferente.

Hoje, o emprego de comandos num ambiente baseado em menus, proporciona

[...] ao usuário uma lista de comandos dentre os quais ele pode selecionar os que forem apropriados. Isso ainda exige do usuário que ele possua alguma noção do efeito que terá a aplicação de determinado comando, porém elimina a necessidade de ele ter de lembrar a forma exata de um comando num dado sistema de recuperação da informação (ROWLEY, 2002, p. 168).

Na Web, portanto, a informação foi sendo adicionada em escala exponencial e máquinas de busca foram priorizando, classificando e manipulando dados para fazer uma recuperação da informação sempre mais sofisticada e fácil.

Deste modo, conforme afirma Mercado (1999, p. 261, tradução nossa), “a interface World Wide Web é a nova escolha para bases de dados online e outras informações, e os

---

<sup>11</sup> [...] the Internet developed from the text-based to the graphics interface of the World Wide Web. Its information became totally accessible with the use of graphics software such as Mosaic and Netscape, which organized the information into an attractive and easily searchable form.

principais sistemas de busca e vendedores de CD-ROM estão fazendo do acesso a Web sua prioridade”<sup>12</sup>.

Após esta breve explanação vê-se, então, que o que vem ocorrendo, devido aos constantes avanços das tecnologias de informação e comunicação, é que cada vez mais,

a capacidade de armazenamento em bancos de dados, assim como sua utilização, vem crescendo na mesma proporção dos avanços em novas tecnologias de informação e comunicação, [e] a atividade de extrair informações relevantes, por conseguinte, está se tornando bastante complexa (TARAPANOFF, et. al. 2001, p. 20).

Ou seja, a intrincada atividade de extrair informações pertinentes à resolução de um problema qualquer de dentro de um sistema de informação, leva a pensar e a pesquisar suas implicações e os processos de interação entre humanos e máquinas. É o que aborda o próximo capítulo: Recuperação da Informação em Bases de Dados.

---

<sup>12</sup> The World Wide Web interface is now the choice for online databases and other information, and major search systems and CD-ROM vendors are making Web access their priority.

## CAPÍTULO 3

Metade do conhecimento consiste em saber onde encontrá-lo<sup>13</sup>.

[...] você não precisa saber tudo, basta saber [como] encontrar.  
Wurman<sup>14</sup>

### RECUPERAÇÃO DA INFORMAÇÃO EM BASES DE DADOS.

A complexidade dos sistemas de recuperação da informação em bases de dados envolve-se, basicamente, em quatro etapas sendo elas: a representação descritiva, a indexação (também chamada de representação de conteúdo ou representação temática), o armazenamento e o processo de recuperação propriamente dito (GRIGÓRIO, 2002; ROWLEY, 2002). Todas essas etapas, conforme Grigório (2002, p. 43, grifo nosso),

[...] tem como objetivo a recuperação de um determinado documento ou conjunto, em meio a uma gama de documentos, ou seja, no processo de inclusão em um índice utilizam-se diversos instrumentos e técnicas para representar cada documento como único e, no sentido oposto, quais as estratégias de busca possíveis para a sua recuperação. Duas fases em especial fazem parte deste processo: a **representação descritiva** do documento, também chamada comumente de **catalogação**, que identifica

---

<sup>13</sup> Trata-se de uma alusão à tão citada afirmação feita pelo notável Dr. Samuel Johnson, há mais de 200 anos: ‘o conhecimento é de duas espécies. Podemos conhecer nós mesmos um assunto ou saber onde podemos encontrar informações a respeito’ (GROGAN, 1995, p. 7).

<sup>14</sup> WURMAN, Richard Saul. **Ansiedade de informação**. Tradução: Virgílio Greire. São Paulo: Cultura Editores Associados, 1991. (Tradução do original: Information Anxiety)



autores, títulos, fontes de publicação, e outros elementos bibliográficos; e, a **indexação de assuntos** que identifica os assuntos de que trata o documento.

Neste sentido, pode-se argumentar que a recuperação da informação está preocupada com os processos que envolvem a **representação** (tanto descritiva, quanto temática), o armazenamento, a busca, e a obtenção da informação que é pertinente para a demanda de um usuário (INGWERSEN, 1992). Entretanto como destaca Foster et al. (2002), o uso de sistemas de recuperação da informação é uma possível potencialização nos processos de busca da informação.

### 3.1 Representação Descritiva ou Catalogação

A representação descritiva ou catalográfica compõem-se, basicamente, em uma forma de descrever um item documentário de forma a torná-lo único e identificável. Comumente chamada de Catalogação, segundo Flamino e Santos (2004, p. 115),

[...] consiste em representar itens documentários de forma clara sucinta e padronizada, individualizando-os e tornando-os acessíveis por várias formas de acesso para que o usuário possa recuperá-lo com rapidez e objetividade. A catalogação tem por objetivo a recuperação da informação, isto é, dizer onde está o documento.

Ou seja, “a catalogação é uma das formas encontradas para organizar e ao mesmo tempo recuperar o conhecimento produzido” (FLAMINO; SANTOS, 2004, p. 115).

Assim, nas bases de dados, a Catalogação permite que itens documentários sejam recuperados, principalmente, pelo:

- Título;
- Indicação de Responsabilidade (responsáveis pelo conteúdo intelectual – Autor, Organizador, Coordenador, etc);
- Série;
- Detalhes específicos (tipos de material/documentos: periódicos, iconográficos, cartográficos, audiovisuais, jurídicos, etc);
- Entre outros.

Para se catalogar itens documentários, há a necessidade da utilização de ferramentas que otimizem a padronização, de forma que seja possível o intercâmbio de registros (com o intuito de reduzir tempo e custos com a catalogação) e, principalmente, a recuperação do item documentário que, conseqüentemente, suporta a informação que o usuário deseja ou necessita.

### **3.2 Representação temática ou Indexação e as Linguagens de indexação**

Os processos de operacionalização da recuperação da informação dependem também, em alto grau, das etapas de indexação e armazenamento.

De acordo com Lancaster (1993, p.4),

[...] são muitos os fatores que determinam se uma busca numa base de dados é ou não bem sucedida. Entre tais fatores encontra-se a abrangência da base de dados, sua política de indexação, suas regras de indexação, sua política e regras para redação de resumos, a qualidade do vocabulário empregado na indexação, [...] e por aí a fora.

A indexação é uma das formas de descrição de conteúdo. É a operação pela qual escolhe-se os termos mais apropriados para descrever o conteúdo de um documento (MENOU; GUINCHAT, 1994). Para Ribeiro e Santos (2003, p. 122) indexação “é a ação de descrever e identificar um documento de acordo com seu assunto” [...], isto é, “uma operação que consiste em extrair os elementos que caracterizam o conteúdo do documento para se obter uma síntese mediante a atribuição de um ou mais termos, com a finalidade de recuperar a informação”.

Os fatores que exercem influência no desempenho de sistemas de recuperação da informação e que estão relacionados com a indexação podem ser categorizados, segundo Lancaster (1993, p. 22) em:

1. Política de indexação;
2. Exatidão da indexação (Análise conceitual; tradução).

A decisão **política** fundamental refere-se à exaustividade da indexação, ou seja, ao número de termos adotados para abranger o conteúdo temático do documento do modo mais completo possível. A outra seria a especificidade, isto é, um assunto ou conceito de um documento deve ser indexado sob o termo mais específico que o represente. Juntas, representam as duas dimensões da indexação de um documento. De outro modo (Lancaster, 1993) define que,

- **Exaustividade** – é a extensão com que a diversidade do conteúdo temático, que é objeto de análise de um dado documento, se acha abrangida numa representação desse documento. Corresponderá, aproximadamente, ao número de termos de indexação atribuídos ou alguma outra medida do número de pontos de acesso criados (por

exemplo, a extensão de um resumo). A exaustividade ideal é inteiramente dependente da demanda.

- **Especificidade** – é o nível de detalhe com que um tópico examinado num documento é abrangido numa representação desse documento. O nível de especificidade possível é determinado pelas características do vocabulário empregado na indexação, embora o indexador (provavelmente de modo equivocado) possa representar um tópico de forma menos específica do que aquela que o vocabulário permite.

Já do ponto de vista da exatidão o processo de indexação, segundo Fujita (2003) e Grigório (2002) consiste na:

- **Análise conceitual** – que implica em decidir do que se trata um documento, isto é, qual o seu **assunto**<sup>15</sup>. Ou seja, no processo analítico realiza-se a compreensão do texto como um todo, identificando e selecionando **conceitos**<sup>16</sup> válidos para a indexação.
- **Tradução** – que envolve a conversão da análise conceitual de um documento num determinado conjunto de termos de indexação, geralmente extraídos de alguma forma de vocabulário controlado. Quer dizer, a representação de conceitos por meio de termos obtidos de uma

---

<sup>15</sup> Assunto: Tema representado num documento por um conceito ou combinação de conceitos (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992, p.1).

<sup>16</sup> Conceito: Qualquer unidade de pensamento. O conceito pode ter o seu conteúdo semântico reexpresso pela combinação de outros conceitos, que podem variar de uma língua ou de uma cultura para outra (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 1992, p.1).

linguagem de indexação, de forma a traduzir os conceitos nos termos de linguagem de indexação.

Isto é, a indexação compõe-se em representar o conteúdo do documento através de conceitos retirados do próprio texto e que podem ser atribuídos através de uma linguagem de indexação como, por exemplo, os Tesouros.

Para Rowley (2002, p.162) a,

Indexação é o processo de atribuir termos ou códigos de indexação a um registro ou documento, termos ou códigos esses que serão úteis posteriormente na recuperação do documento ou registro. A atribuição dos termos de indexação pode ser intelectual (ou seja, realizada por um ser humano) ou feita automaticamente pelo computador, que, no entanto, somente pode selecionar termos de indexação de acordo com um conjunto de instruções. A seleção dependerá da ocorrência das palavras. Os termos de indexação passíveis de serem atribuídos serão extraídos de uma lista-padrão ou tesouro instalado no computador, com base na ocorrência de palavras num registro ou documento. Alternativamente, os computadores podem ser também utilizados para pôr em ordem termos de indexação atribuídos por seres humanos. Esses termos serão selecionados e atribuídos por indexadores com base no julgamento subjetivo que fazem acerca do conteúdo do documento, ou escolhem termos que tenham a probabilidade de virem a ser procurados por um usuário no futuro. Os termos podem ser extraídos de uma lista controlada de termos ou poderão ser livres de controle. Muitos sistemas incorporam elementos tanto de linguagens de indexação controladas quanto não controladas. O computador funciona como um burro de carga confiável que organizará os verbetes do índice em

ordem alfabética para sua apresentação na tela ou para a impressão de um índice.

Uma linguagem de indexação é uma linguagem usada para descrever assuntos ou outros aspectos de informação ou documentos em um índice (ROWLEY, 1988). As duas categorias principais são como segue.

A **linguagem controlada** ou o vocabulário controlado é uma linguagem de indexação onde os sinônimos, homógrafos, plural e singular são controlados de forma a facilitar a busca (RIBEIRO; SANTOS, 2003). Conforme Lopes (2002a, p.47, grifo nosso) a linguagem controlada ou o vocabulário controlado,

pode ser definido como um conjunto de termos organizados de forma hierarquizada e/ou alfabética, com o objetivo de possibilitar a recuperação de **informações temáticas**, reduzindo substancialmente a diversidade de terminologia.

Os **tesauros**, que são produtos do uso da linguagem controlada, para Silveira (2002) “constituem uma base de conhecimento sobre a terminologia utilizada em uma dada área do conhecimento” com o “objetivo de representar conceitos e especificar seus relacionamentos” ou ainda para Ribeiro e Santos (2003, p. 236) são um

Conjunto de conceitos ordenados, de modo claro e livre de ambigüidade, a partir do estabelecimento de relações entre os mesmos e pode ser definido segundo sua função ou estrutura. Do ponto de vista de sua função, é um instrumento de controle terminológico adotado por sistemas e/ou centros de informação e bibliotecas com o objetivo de tornar a indexação do conteúdo

temático de documentos textuais/bibliográficos mais consistente e, conseqüentemente, garantir maior precisão na recuperação de informações. Quanto a sua estrutura, é um vocabulário controlado e dinâmico de termos que têm entre si relações semânticas e genéricas, e aplica-se a uma área particular do conhecimento.

Já a **linguagem natural** ou vocabulário livre seria como o “sinônimo do discurso comum, isto é, a linguagem usada habitualmente na fala e na escrita sendo que, **nas bases de dados**, os **termos** do **título** e **resumo** representam a linguagem natural” (LOPES, 2002a, p. 48, grifo nosso). Ou seja, é a

Linguagem de indexação onde os termos são extraídos do documento que está sendo indexado. Não existe o risco de ocorrer distorções das palavras do autor, no entanto não há controle de sinônimos e ambigüidades exigindo grande esforço no estágio da busca [...] (RIBEIRO; SANTOS, 2003, p. 147).

Uma vez que “não se trata efetivamente de linguagens precisas ou invariáveis em si mesmas” (ROWLEY, 2002, p.170), a linguagem natural é a linguagem do discurso técnico-científico, e, no contexto da recuperação da informação, refere-se às palavras que aparecem nos textos impressos considerando-se como seu sinônimo a expressão “texto livre” (LANCASTER, 1993).

“Nas bases de dados, os campos de título e resumo registram os termos da Linguagem Natural enquanto os campos de descritores, termos de indexação ou identificadores registram os termos da Linguagem Controlada” (LOPES, 2002a, p. 42).

Ambas trazem suas vantagens e desvantagens que podem ser vistas nas figuras 8 e 9 a seguir:

#### **LINGUAGEM CONTROLADA:**

<b>VANTAGENS</b>	<b>DESVANTAGENS</b>
Controle total do vocabulário de indexação, minimizando os problemas de comunicação entre indexadores e usuários.	Custos: a produção e manutenção da base de dados terão despesas maiores com a equipe de indexadores. Será necessário ainda manter pessoal especializado na atualização do thesaurus.
Com o uso de um thesaurus e suas respectivas notas de escopo, os indexadores podem assinalar mais corretamente os conceitos dos documentos.	O vocabulário controlado poderá não refletir adequadamente os objetivos do produtor da base, caso esteja desatualizado.
Se bem constituído, o vocabulário controlado poderá oferecer alta recuperação e relevância e, também ampliar a confiança do usuário diante de um possível resultado negativo.	Um vocabulário controlado poderá distanciar dos conceitos adequados para a representação das necessidades de informação dos usuários.
As relações hierárquicas e as remissivas do vocabulário controlado auxiliam tanto o indexador, quanto o usuário na identificação de conceitos relacionados.	Necessidades de treinamento no uso dos vocabulários controlados tanto para os intermediários, quanto para os usuários finais.
Redução no tempo de consulta á base, pois a estratégia de busca será mais bem elaborada com o uso do thesaurus.	Desatualização do vocabulário controlado poderá conduzir a falsos resultados.

**Figura 4:** Vocabulário controlado: vantagens e desvantagens

**Fonte:** LOPES, 2002, p. 47



### LINGUAGEM NATURAL:

VANTAGENS	DESVANTAGENS
Permite o imediato registro da informação em uma base de dados, sem necessidade de consulta a uma linguagem de controle.	Os usuários da informação, no processo de busca, precisam fazer um esforço intelectual maior para identificar os sinônimos, as grafias alternativas, os homônimos etc.
Processo de busca é facilitado com a ausência de treinamentos específicos no uso de uma linguagem de controle.	Haverá alta incidência de respostas negativas ou de relações incorretas entre os termos usados na busca (por ausência de padronização).
Termos de entrada de dados são extraídos diretamente dos documentos que vão constituir a base de dados.	Custos de acesso tendem a aumentar com a entrada de termos de busca aleatórios.
Temas específicos citados nos documentos podem ser encontrados.	Uma estratégia de busca que arrole todos os principais conceitos e seus sinônimos deve ser elaborada para cada base de dados (ex: nomes comerciais de substâncias químicas não ocorrem no Chemical Abstracts).
Elimina os conflitos de comunicação entre indexadores e os usuários, pois ambos terão acesso aos mesmos termos.	Perda de confiança do usuário em uma possível resposta negativa.

**Figura 5:** Linguagem Natural: vantagens e desvantagens

**Fonte:** LOPES, 2002, p. 48

Sobre o uso da linguagem controlada, vale salientar que,

[...] na comunicação lingüística as palavras se clarificam entre si segundo o contexto, nas buscas documentais esta ambigüidade é um obstáculo que conduz o usuário a encontrar documentos distintos dos buscados. Isto obriga a estabelecer os termos de indexação de forma mais unívoca

possível, para evitar respostas parasitas ou ruído documental nas recuperações. As linguagens controladas buscam uma aproximação entre os conceitos documentais e os lingüísticos que melhore a análise da informação armazenada nas bases de dados documentais, para sua exata recuperação (LÓPES ALONSO, 1999, p. 283, tradução nossa)<sup>17</sup>.

Quanto ao uso da linguagem natural, López Alonso (1999) afirma que embora há uma melhora na relevância nas recuperações de documentos nos sistemas de gestão de informação, a inconsistência na representação da informação aparece relacionada com a indeterminação da linguagem natural para a representação do conhecimento.

Logo, vemos que, a recuperação com linguagem natural produz maior revocação e uma menor precisão que a realizada com um vocabulário controlado. Mas uma menor revocação e maior precisão inerente à recuperação realizada com as linguagens controladas, depende do fato intrínseco que as caracteriza como sendo uma redução na quantidade de termos usados para a busca. E, subsidiariamente, da qualidade dos vocabulários controlados utilizados, por exemplo: do número de termos nomeados para cada descritor, de sua capacidade para evitar os homônimos, etc (SOERGEL, 1994). Ou, como Lancaster (1993, p.4) exprime que “uma característica lamentável, inerente à recuperação da informação, é uma melhoria da revocação em geral implicar uma perda de precisão e vice-versa”.

---

<sup>17</sup> [...] en la comunicación lingüística las palabras se clarifican entre sí según el contexto, en las búsquedas documentales esta ambigüedad es un obstáculo que conduce al usuario a encontrar documentos distintos de los buscados. Esto obliga a establecer los términos de indización de la forma más unívoca posible, para evitar las respuestas “parásitas” o ruido documental en las recuperaciones. Los Lenguajes Controlados buscan un acercamiento entre los conceptos documentales y los lingüísticos que mejore el análisis de la información almacenada en las Base de Datos Documentales, para su exata recuperación.

Precisão e revocação são medidas do desempenho para buscas realizadas numa base de dados. Revocação é definido como a proporção de material pertinente recuperado, enquanto Precisão é a proporção de material recuperado que é pertinente (MA, 2002).

Segundo Lancaster (1993) e González Olivares e García Gómez (2001) em outras palavras,

- **Precisão** – a extensão com que os itens recuperados durante uma busca numa base de dados são considerados relevantes ou pertinentes. Uma busca que alcance uma precisão alta será aquela em que a maioria dos itens recuperados, se não todos, forem considerados relevantes ou pertinentes. O coeficiente de precisão (**P**) – uma medida da extensão com que se alcança a precisão – é o número de **documentos relevantes recuperados (drr)** dividido pela **soma dos documentos relevantes recuperados (drr)** e os **documentos não relevantes recuperados (dnrr)** ( $P = drr / (drr + dnrr)$ ).
- **Revocação** – A extensão com que todos os itens numa base de dados que são considerados relevantes ou pertinentes são recuperados durante uma busca nessa base de dados. Uma busca com “revocação alta” será aquela em que a maioria dos itens relevante (pertinentes) for recuperada. O coeficiente de revocação (**R**) – uma medida da extensão com que ocorre a recuperação de itens relevantes – é o número de **documentos relevantes recuperados (drr)** dividido pelo número **total de documentos relevantes (Tdr)** (pertinentes) existentes na base de dados ( $R = drr / Tdr$ ).

Em outras palavras, **Revocação** mede a proporção de itens relevantes na coleção que foi de fato recuperado pelo sistema, e é então uma medida de eficiência da habilidade de recuperação do sistema. **Precisão**, por outro lado, pega o número total de documentos recuperados, e mede a porcentagem desses documentos recuperados que são de fato relevantes (MA, 2002). Porém, o conceito de Revocação pode ser também aqueles “documentos selecionados e considerados não relevantes para a pesquisa” (RIBEIRO; SANTOS, 2003, p. 213).

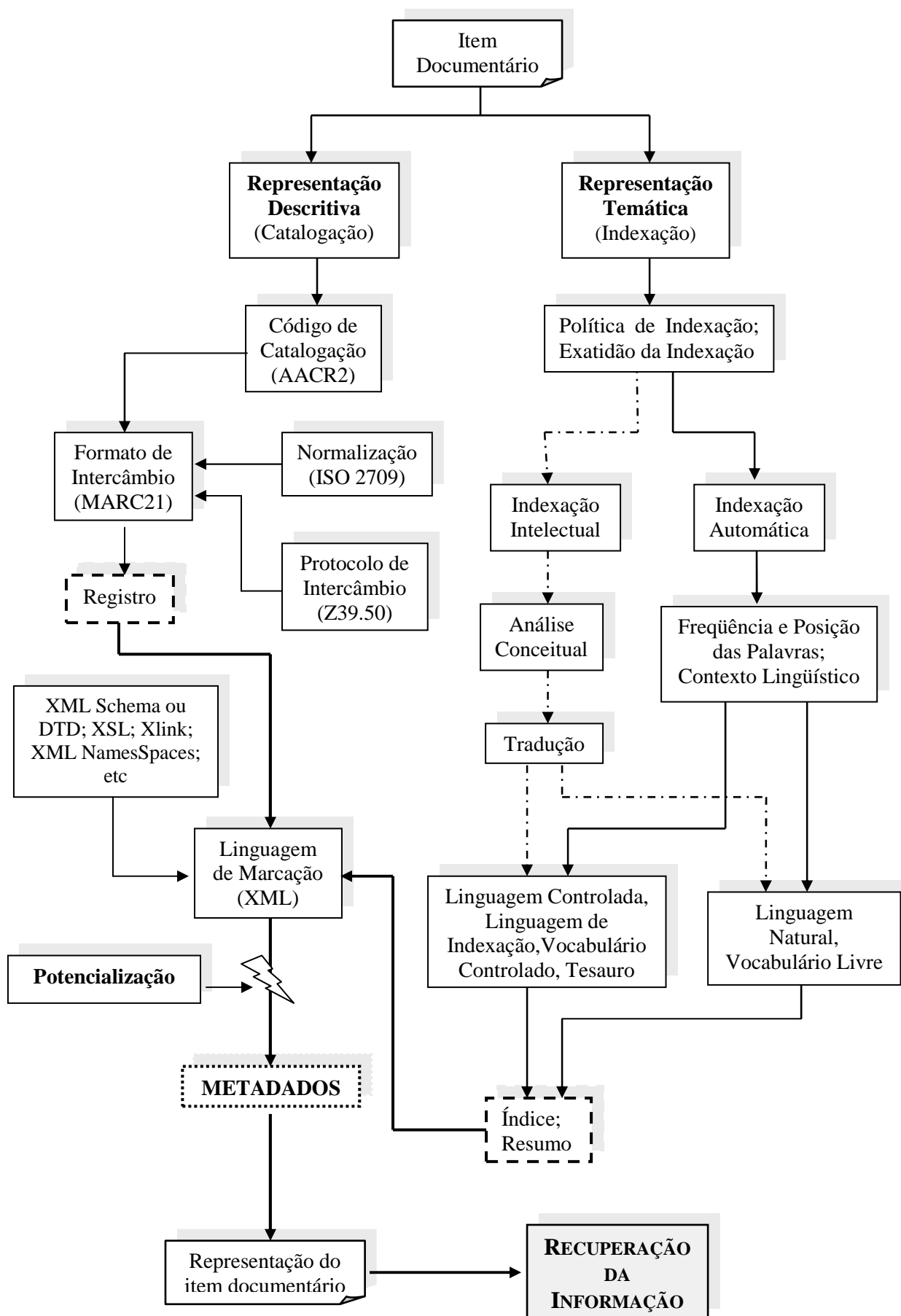
De qualquer forma, muitas bases e bancos de dados oferecem, cada vez mais, a possibilidade de busca utilizando tanto a linguagem natural como a linguagem controlada simultaneamente (comumente chamados de sistemas híbridos<sup>18</sup>), recomenda-se que uma estratégia de termos em linguagem natural e em linguagem controlada deva ser adotada, a fim de serem obtidos resultados satisfatórios na recuperação. Muddamalle (1998, p. 887, tradução nossa) salienta que “a linguagem controlada e a linguagem natural não podem mais ser tratadas como técnicas de buscas separadas, mas devem sempre ser tratadas em conjunto, como uma combinação ideal para ampliar os resultados das buscas de informação”.

### 3.3 Metadados

Um exemplo do fluxo da representação do item documentário pode ser apresentado conforme a figura que se segue:

---

<sup>18</sup> O termo *híbrido* é empregado para designar qualquer sistema que funcione com uma combinação de termos controlados e linguagem natural, inclusive aqueles em que ambos os conjuntos de termos são atribuídos por indexadores humanos e aqueles em que uma base de dados pode ser consultada mediante uma combinação de termos controlados atribuídos por seres humanos e palavras que ocorram nos títulos, resumos ou texto completo (LANCASTER, 1993, p. 223).



**FIGURA 6 :** Exemplo de fluxo da representação do item documentário para a recuperação da informação

Na figura anterior, pode-se ver que tanto o processo da representação descritiva quanto o processo da representação temática, utilizam-se de ferramentas e técnicas (de forma cada vez mais integrada) para a geração de produtos que possibilitarão que um documento específico seja recuperado. Estes produtos podem ser definidos como Metadados, que podem ser mais comumente definidos como dados sobre outros dados (SOUZA; VENDRUSCULO; MELO, 2000). Segundo, Perez (2000, p. 26),

A palavra "metadado" possui a mesma origem grega da palavra "metamorfismo", ou seja, o termo "meta" significa "mudança" de tal forma que metadado descreve a origem e as mudanças do(s) dado(s) [...]. Uma legenda de um mapa é claramente um metadado. Ela contém informação sobre o autor do mapa, data de publicação, tipo, descrição, referências espaciais, escala e precisão, dentre outras. Ou seja, um metadado é esse tipo de informação descritiva aplicada a um objeto.

Para Milstead e Feldman (1999, p. 2, tradução nossa) “[...] a informação bibliográfica padrão, os índices, os termos indexadores, e os “abstracts” são todos representantes do material original, portanto metadados”<sup>19</sup>. Isto é, pode-se dizer que são

[...] um conjunto de dados usados para descrever e representar objetos informacionais, assim como, descrever o seu conteúdo e sua localização e desta forma facilitar o acesso e recuperação de informações seja na Internet ou não, pois o termo metadado se refere a dado em sentido amplo e não somente para fontes eletrônicas. O resumo de um texto ou as suas palavras-

---

<sup>19</sup> [...] standard bibliographic information, summaries, indexing terms, and abstracts are all surrogates for the original material, hence metadata.

chave, por exemplo, representam o texto original, são dados sobre outros dados, portanto considerados metadados (FLAMINO, 2003, p. 67).

Assim, com o uso de vocabulários controlados, a limitação de buscas aos descritores, identificadores, autores, títulos, ou campos de fontes, tem sido capazes de melhorar a precisão das buscas por muitos anos. Com a utilização do modelo booleano, os metadados tornaram possível encontrar a informação sem o excessivo aparecimento de disparates estranhos (MILSTEAD; FELDMAN, 1999).

Entretanto,

[...] Utilizar os metadados, procedimento que no campo bibliotecário se tem utilizado mais para a atividade descritiva que qualquer outra da gestão de informação, mas que podem ter distintos usos potenciais, todos eles estão direcionados ao descobrimento e recuperação de objetos de informação da forma mais simples e eficaz possível. Neste sentido se tem desenvolvido alguns formatos para metadados específicos no campo bibliotecário, que parece ser que estão dando bons resultados, embora não estão totalmente consolidados, devido a fase de experimentação em que se encontram [...] (GARCÍA MARTÍNEZ, 2001, p. 46, tradução nossa)<sup>20</sup>.

---

<sup>20</sup> [...] utilizar los metadatos, procedimiento que el campo bibliotecario se ha utilizado más para la actividad descriptiva que cualquier otra de la gestión de información, pero que puede tener distintos usos potenciales, todos ellos encaminados al descubrimiento y recuperación de objetos de información de la forma más sencilla y eficaz posible. En este sentido se han desarrollado algunos formatos de matadatos específicos en el campo bibliotecario, que parece ser que están dando buenos resultadosm aunque no están totalmente consolidados, debido a la fase de experiementación en que se encuentran [...].

Pode-se ver na literatura especializada, muitos exemplos da aplicação dos metadados para a representação da informação, principalmente com as iniciativas que a XML vem trazendo para a **potencialização** da representação do item documentário para a otimização dos processos de recuperação da informação.

### 3.3.1 Formato de Intercâmbio MARC 21

A mais comumente ferramenta conhecida e utilizada por diversas unidades de informação, mais notavelmente em bibliotecas, para a representação é o formato de intercâmbio MARC 21, um registro catalográfico legível por computador.

O formato de intercâmbio MARC 21, segundo Balby (1995), compõe-se de três partes:

- **Formato de entrada** – usado pelo catalogador na entrada dos dados;
- **Formato de armazenamento interno** – refere-se à alocação dos dados na memória do computador;
- **Formato de intercâmbio** – utilizado para a transferência e comunicação de dados bibliográficos entre sistemas de computadores diferentes e/ou remotos, gerado a partir do formato de armazenamento interno.

E a estrutura de um registro MARC 21 é composta de três elementos Básicos: segundo Flamino e Santos (2004), Furrie (2000), MARC21 (2002) são:

- **Líder** – contém os dados que fornecem as informações para o processamento do registro pelo próprio sistema operacional;
- **Diretório** – informa ao sistema operacional que campos estão no registro, onde estão localizados e o tamanho de cada campo;



- **Campos** – o registro é dividido em campos (marcados por etiquetas) e subcampos. Para cada indicação (autor, título, imprensa, etc) há um campo específico.

Para a descrição bibliográfica, o sistema adota o Código de Catalogação Anglo-Americano, 2ª edição (**AACR2**, na sigla em Inglês). Ele padroniza a entrada dos dados no formato MARC 21, definindo os campos para determinado item documentário e a entrada dos seus respectivos dados. A norma internacional **ISO 2709** aplica-se para a padronização, especificando os requisitos para o formato de intercâmbio de registros bibliográficos. E para a comunicação entre diferentes sistemas de computadores possibilitando a recuperação e o intercâmbio de dados, o formato MARC 21 utiliza-se do Protocolo de Comunicação **Z39.50** (FLAMINO; SANTOS, 2004).

Porém, com a dinamização dos recursos eletrônicos e digitais, devido aos desenvolvimentos tecnológicos e ao intenso uso da Internet, surgiu a necessidade de ferramentas mais apropriadas para este ambiente, principalmente no que diz respeito ao gerenciamento, tratamento, armazenamento, recuperação e transmissão de informações via Internet, como a XML.

### 3.3.2 EXtensible Markup Language – XML

A **XML** (eXtensible Markup Language), vem com a proposta de **potencializar** os recursos disponíveis na WEB, inclusive na representação descritiva e temática de itens documentários. Ela permite tornar possível a marcação dos documentos, considerando a semântica da informação, isto é, cada instituição pode produzir o conjunto de campos que

melhor represente os elementos que compõem os seus documentos (MARTÍNEZ GONZÁLEZ, 2000).

Segundo Furgeri (1999, p. 26) “Um documento XML é composto, basicamente, de três elementos distintos”:

- Conteúdo dos dados: são as informações entre as *Tags*;
- Estrutura: a organização dos elementos dentro do documento (pode possuir a estrutura de um memorando, contrato, receita etc.);
- Apresentação: é a forma com que as informações são apresentadas ao leitor do documento.

Com inúmeras vantagens a XML possibilita, entre outras (FLAMINO; SANTOS, 2004; ORTIZ-REPISO JIMÉNES, 1999; FURGERI, 1999):

- Pode ser usada para criar registros catalográficos e para criar os textos eletrônicos do catálogo;
- O armazenamento da informação e sua descrição num mesmo local ou seja, os documentos XML são essencialmente banco de dados de informações, podendo ser consultados e processados como qualquer banco de dados;
- Não há limitações para os tamanhos dos documentos;
- Sua estrutura permite a indexação e a recuperação do documento na íntegra, ou as partes do documento, mediante busca booleana, adjacência, proximidade, ranking de relevância e outros;
- O documento de mesmo conteúdo pode ser visualizado de diferentes formas, para diferentes usuários, através da utilização de folhas de estilo;

- Permite alto grau de automação para agentes de software (agentes inteligentes);
- A XML possibilita a entrega de qualquer tipo de dados estruturado por toda a WEB. Aplicações XML padronizadas possibilitarão que diferentes aplicativos trabalhem em conjunto, significando uma maior **Interoperabilidade**, ou seja, a compatibilidade entre sistemas e plataformas heterogêneos.

Para Furgeri (1999, p. 26),

A XML torna o documento mais “inteligente” criando uma estrutura que mostra claramente qual é o significado das informações, pois elas estão agrupadas, e não dispersas pelo documento. Essas novas etiquetas são criadas a critério do desenvolvedor, [...] Sendo assim as etiquetas são criadas conforme as necessidades dos autores do documento.

É importante ressaltar que “são as tecnologias associadas a XML que ampliam seu potencial” (FLAMINO; SANTOS, 2004, p. 128), entre algumas delas, destacam-se (ALVARENGA; SOUSA, 2004; FLAMINO, 2003):

- **XML Schema**<sup>21</sup> – expressa vocabulários compartilhados e permite que máquinas executem regras elaboradas por pessoas.

---

<sup>21</sup> A linguagem XML Schema, apesar de ter a mesma função da DTD, possui muitas características que a torna mais poderosa (e mais completa) que a DTD. A DTD (Document Type Definition) é um arquivo do tipo texto onde estão definidas as tags [etiquetas], a ordem em que elas devem aparecer no documento XML e sua obrigatoriedade (FERNEDA, 2003).

Provê meios para definir a estrutura, conteúdo e semântica de documentos XML;

- **XSL** – usada para converter um documento XML em diferentes formatos de exibição e para diferentes dispositivos;
- **Xlink** – permite inserir elementos em documentos XML para criar e descrever ligações entre recursos (hyperlinks), mais potentes com características de extensibilidade, inclusão de informação acerca da semântica e da relação, links bidirecionais, links múltiplos, links fora dos documentos, ligar fragmentos de documentos, manipulação como qualquer outro dado XML.
- **XML Namespaces** – o desenvolvedor pode criar os seus próprios elementos com os nomes que quiser. São os nomes dos elementos relacionados no interior das etiquetas, um dos trunfos da XML.

Desta forma, assim como os negócios e a indústria, a representação e organização da informação na era digital tiveram mudanças dramáticas em quase todas as fases do seu processo. As mudanças incluem não somente os métodos e a tecnologia usada na criação de registros para publicações, mas também nos padrões que são essenciais para o sucesso e a efetividade das ferramentas de busca e recuperação da informação. Hoje os catálogos das bibliotecas estão longe de ser somente uma ferramenta para sua própria coleção e para o uso dos visitantes das bibliotecas; Estes se tornaram um nó na rede os quais os usuários podem visitar de qualquer lugar do mundo através de um computador conectado à Internet (QIN, 2000).

Percebe-se, então, que neste contexto “a união da estrutura de representações bibliográficas e catalográficas MARC e a linguagem XML serão extremamente valiosas [...]”

(FLAMINO; SANTOS, 2004, p.132) no que diz respeito à busca e a recuperação de informações.

### **3.4 Armazenamento**

Muitas bases e bancos de dados que disponibilizam serviços de busca on-line, empregam sistemas de gerenciamento de documentos utilizam quase sempre o método de arquivo invertido, pois, é útil nas buscas em bases de dados textuais complexas. Segundo Menou e Guinchat (1994, p. 319) “a pesquisa em arquivo invertido [...] consiste em comparar cada um dos pontos de acesso que figuram nas questões, com os pontos do arquivo invertido a partir das informações dos registros bibliográficos”. Esta “pesquisa”, que compara os pontos de acesso digitados pelo usuário com os pontos do arquivo invertido, é feito pelo sistema de busca.

Um arquivo invertido, também conhecido como índice invertido, é um índice ordenado de palavras-chaves, com cada palavra-chave contendo encadeamentos para os documentos que as contém (BATISTA, 2004). Segundo Ribeiro e Santos (2003, p. 22),

Este termo é utilizado porque todos os termos indexados numa base de dados estão armazenados em um arquivo invertido gerado pelo programa. Às palavras deste arquivo estão associados os números de registros que as contêm. Assim, ao invés de buscar no texto normal, o programa busca nesse arquivo que é bem menor e, quando encontra a palavra, recupera os registros associados a ela.

Neste sentido, quando se utiliza a linguagem natural para a recuperação, esta se armazenando, como arquivo invertido, pelo menos tradicionalmente, as palavras mais

significativas dos “campos” título e resumo do documento que são bem menores que o documento completo (ROWLEY, 2002). Ou ainda, os mecanismos se baseiam na análise estatística da frequência relativa de ocorrência dos termos no **texto integral**, e a indexação, baseada em linguagem natural, (que pode ser realizada por um indexador humano, mas que, cada vez mais, é realizada automaticamente pelo computador) é feita com cada termo presente em todo o documento, exceto os de uma lista de termos proibidos, de extensão limitada, com palavras muito comuns (como artigos e preposições); ou ainda, apenas com os termos que se encontrem num tesouro armazenado na memória do computador (ROWLEY, 2002). Bases de dados como a Emerald-Library<sup>22</sup> dão a possibilidade de busca no texto completo. Mas utilizar apenas este recurso, certamente trará resultados com muito pouca precisão, visto que as palavras encontradas no texto completo são muita menos precisas para representar o documento, do que as encontradas no título ou resumo.

### **3.5 Etapas do processo de recuperação da informação**

Há numerosos estudos, principalmente na literatura especializada estrangeira, que descrevem os complicados processos estratégicos de busca e recuperação que deveriam ser empreendidos ao se explorar os domínios de informação (DEBOWSKI, 2001).

Lopes (2002b, p. 65), destaca que

No planejamento da estratégia de busca, algumas ações ou táticas precisam ser criteriosamente relacionadas, principalmente com a seleção de termos e/ou conceitos e com as restrições e/ou limites determinados pelo solicitante entre outras questões.

---

<sup>22</sup> Disponível em: <http://www.emerald-library.com> acesso em: 19 nov. 2004

Para Debowski (2001), são três os estágios para a realização de uma boa busca:

- **Identificação do conceito.** A maioria dos pesquisadores começa com a declaração de um problema o qual necessita ser resolvido. Declarado o problema normalmente inclui-se todos os conceitos fundamentais a ser explorados, permitindo o pesquisador a isolar as palavras chaves que serão submetidas ao processo de busca. Esta primeira fase é importante para assegurar que a busca tenha êxito, sendo que a sequência de busca final necessita incluir todos os conceitos pertinentes para refletir toda a extensão da busca.
- **Seleção.** Como a extensão de busca (número de conceitos) aumentada, o sucesso da busca requer uma pesquisa mais apurada de cada conceito para aumentar a probabilidade da captura da busca. Nestas situações o pesquisador necessita relacionar cada conceito com o seu descritor. O uso de um Tesouro pode ser de grande ajuda provendo os pesquisadores com uma ordem hierárquica de palavras chaves relacionadas, possibilitando a seleção desses termos que estão relacionados com os conceitos do problema. Índices também podem oferecer sugestões para termos de busca e a frequência deles na base de dados, embora eles sejam menos informativos em relação à correlação entre os termos.
- **Seqüência.** O estágio final da busca requer a amalgamação e integração desses vários termos em uma coesiva seqüência ou equação de busca. O

uso de conectores Booleanos sustenta o sucesso da busca em base de dados, com o uso extensivo de convenções tal como “and”, “or”, e “not” para refinar a busca. A equação ou sequência de busca também pode ser formulada pelo uso de parênteses para indicar as relações aninhadas que poderiam ser empregadas. Nesta fase os pesquisadores assumem o entendimento da lógica de busca e as implicações das estruturas de comando.

A autora ainda destaca que,

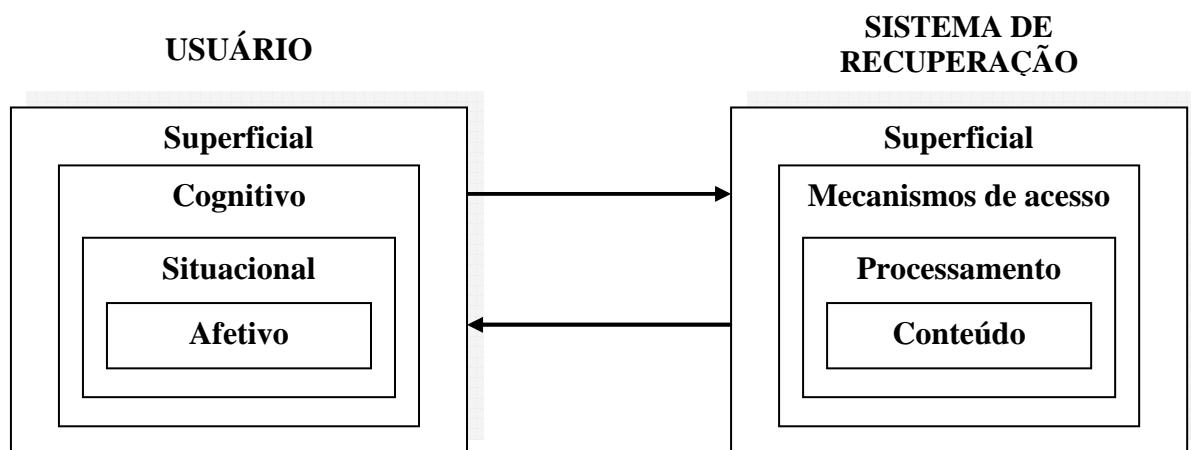
Enquanto este breve esboço das fases de busca parece simples e prontamente aplicado, em buscas eletrônicas raramente é um processo em linha reta. O processo inteiro é cíclico: constantemente opções são avaliadas, selecionadas, postas em prática e revisadas<sup>23</sup> (DEBOWSKI, 2001, p. 373, tradução nossa).

O processo é resultante de interações entre usuário e sistema de recuperação que são entendidas como uma série de episódios no tempo ocorrendo em níveis conectados em série. Cada nível envolve elementos diferentes e/ou processos específicos. Assim, o modelo interativo pode ser representado no esquema a seguir:

---

<sup>23</sup> While this brief outline of the key search stages appears simplistic and readily applied, electronic searching is rarely as straightforward. The entire process is cyclical: options are constantly evaluated, selected, trialled and reviewed.





**Figura 7:** Esquema dos componentes do modelo interativo

**Fonte:** RIBEIRO JÚNIOR, 2003b

No **nível superficial**, autores como Spink e Saracevic (1997), dizem que a interação ocorre em uma série de eventos no tempo, nos quais os usuários interagem por meio de uma interface com o sistema, não somente realizando buscas, mas explorando atributos e informações sobre a interface, sobre os feedbacks gerados, etc. No **nível cognitivo** ocorrem o julgamento e a interpretação dos resultados, onde o usuário interage com os documentos e pode assimilá-los cognitivamente. No **nível situacional**, os usuários interagem com a dada situação ou problema o qual produziu a necessidade de informação que resultou na questão. Os resultados da busca podem ser aplicados na resolução total ou parcial do problema. E, no **nível afetivo**, os usuários interagem com as intenções e motivações, e com sentimentos associados de satisfação, frustração, urgência, etc.

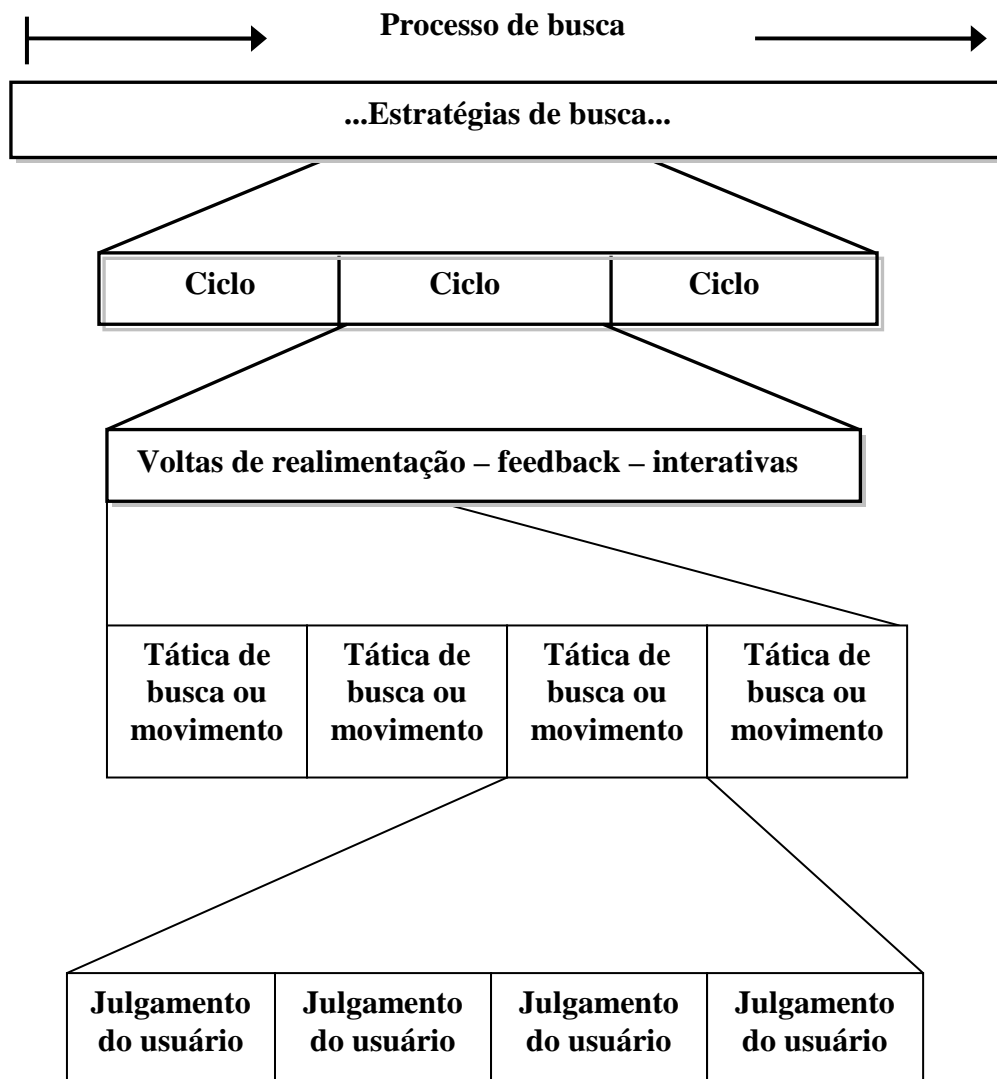
Do lado do sistema, encontramos os mecanismos de acesso aos dados, a interface, o processamento adicional da consulta, se houver, e o conteúdo representado.

Neste sentido, os mesmos autores Spink e Saracevic (1997) afirmam que o discurso do usuário progride através dos episódios, nos quais aspectos dos níveis cognitivos, situacionais e afetivos, na interação, podem mudar o problema ou questão, redefinindo ou mudando o foco, de acordo com um conjunto de satisfações ou frustrações. Este progresso na interação é sinalizado no nível superficial pela mudança de comportamento. Por exemplo, na seleção de novos termos de busca, outros termos são abandonados, táticas de busca são

adaptadas e/ou alteradas por meio da combinação ou substituição de termos, ao contrário do que afirma Rowley (2002, p. 162), quando diz que,

O usuário e as consultas feitas ao sistema em geral não sofrem alterações de um sistema para outro. As necessidades do usuário não se alteram devido à existência de determinados sistemas, ainda que possam tornar-se mais elaboradas à medida que ele adquirir mais experiência com esses sistemas.

Ou seja, certamente, cada estratégia de busca pode consistir em um ou mais ciclos (um ou mais comandos de busca no final da exibição dos itens recuperados). Cada ciclo pode consistir em uma ou mais ocorrências de realimentação – feedback – interativas (entrada/**pergunta** do usuário; saída/**resposta** do sistema de recuperação da informação; interpretação e julgamento do usuário; “entrada” do usuário). Uma “entrada” também pode representar um movimento dentro da estratégia de procura, e pode ser considerada como tática de busca para avançar a busca. E cada movimento consiste em uma “entrada” do usuário ou pergunta requerendo uma “saída” ou resposta do sistema (SPINK, 1997). Veja a figura a seguir:



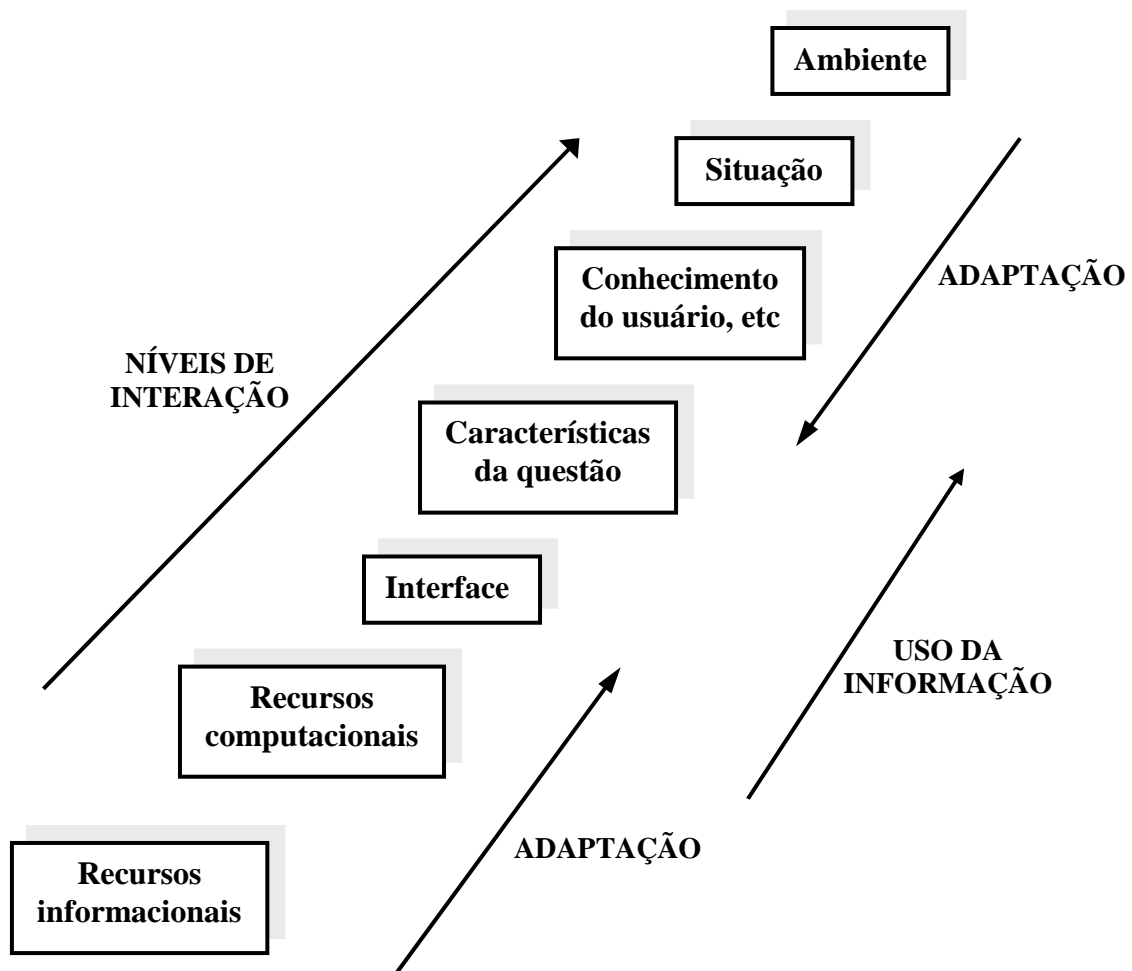
**Figura 8:** Processo de interação da Recuperação da Informação

**Fonte:** SPINK, 1997 (tradução nossa)

Podemos entender também da seguinte maneira. No nível superficial, um usuário interage com um sistema por uma interface emitindo comandos ou perguntas que represente, de algum modo, um problema. No mesmo nível, o sistema responde também com uma meta-informação, ou textos (incluindo imagens, sons, audiovisuais e outros) ou com perguntas formuladas pelo seu próprio projeto para extrair do usuário informação adicional sobre a natureza do problema. No nível cognitivo, o usuário interage com a resposta do sistema, ou com textos obtidos subsequente a interação do sistema, de modo que o capacite para avaliar a utilidade do texto em relação ao problema inicial. No nível situacional, usuários interagem com a determinada situação ou problema a mão que produziu a necessidade de informação,

resultando na questão. E no nível afetivo, os sentimentos associados de satisfação, frustração, urgência interagem com as intenções e motivações dos usuários. Os resultados da busca podem ser aplicados na resolução total ou parcial do problema (SARACEVIC, 1996b).

Podemos ver este processo conforme o esquema abaixo:



**Figura 9:** Processo da Recuperação da Informação

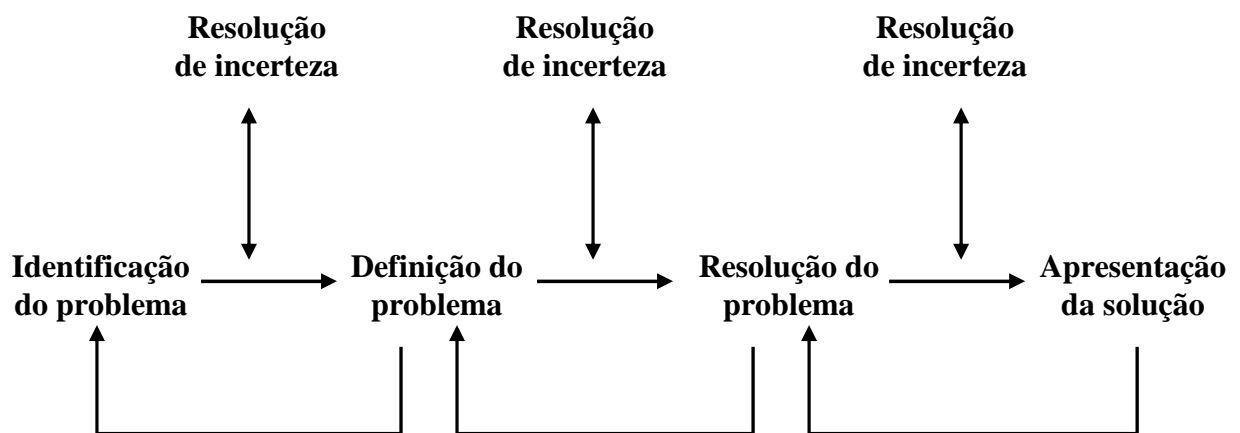
**Fonte:** Saracevic, 1996b (tradução nossa)

Vemos, então, que o indivíduo (pesquisador) move-se da incerteza para a certeza crescente, e que os estágios no processo da resolução do problema são identificáveis e reconhecidos pelo indivíduo. Estes estágios, segundo Wilson (1999), são:

- Identificação do problema (quando uma pessoa está fazendo a pergunta, ‘Que tipo de problema eu tenho?’);

- Definição do problema ('exatamente, qual é a natureza do meu problema?');
- Resolução do problema ('Como eu acho a resposta a meu problema?'), e potencialmente;
- Apresentação da solução (É a resposta ao problema ou, sendo pragmático, ao invés de encontrar uma resolução teoricamente baseada, este estágio é como nós vamos lidar com o problema).

Certamente a transição da identificação do problema para a solução é um trabalho bastante árduo. Cada estágio vê a conversão seguinte com cada vez mais certeza e, onde a incerteza falha ser solucionada, em qualquer dos estágios, isto pode resultar em uma volta de avaliação (feedback) para o estágio prévio, para resolução adicional (WILSON, 1999). Isto pode ser representado na figura seguinte:



**Figura 10:** Resolução do problema no processo de busca de informação

**Fonte:** Wilson, 1999 (tradução nossa)

### 3.6 Estratégias de busca

A possibilidade de acesso aos crescentes e sempre em desenvolvimento sistemas de recuperação de informação, também chamados de bancos de dados e, conseqüentemente,

às suas diversas bases de dados, veio a melhorar significativamente a qualidade das buscas bibliográficas, visto que estas bases proporcionam, cada vez mais, variados pontos de acesso a informação. Segundo Lopes (2002b, p. 60),

Estes sistemas possibilitam o planejamento de estratégias de busca com maior nível de complexidade envolvendo vários conceitos na mesma estratégia; permitem a utilização de busca de palavras apenas dos títulos e resumos dos documentos, isto é, termos da linguagem natural; buscam os termos específicos de linguagem controladas, nos campos de descritor; buscam por autores; por ano de publicação; por títulos de periódicos; por classificação; permitem, também, a busca de conceitos compostos ou simples e a possibilidade de truncagem de raízes de palavras e de substituição de caracteres no meio dos termos, dentre outros recursos de recuperação.

Mas é a partir de uma pergunta, isto é, de um **problema** em questão ou de uma necessidade informacional daquele determinado momento, ou ainda de um tipo de classe especial de estado mental ou psicológico que um indivíduo possui e cujo conteúdo é identificável com um tipo de insatisfação, curiosidade ou desconformidade informativa que,

Normalmente, materializamos e representamos estes estados mentais mediante o enunciado de uma linguagem natural (castelhano, catalão, inglês, [português, etc]). Infelizmente, os sistemas de informação não estão habituados a entender, por assim dizer, essas solicitações ou consultas de informação realizadas a partir dos enunciados de uma linguagem natural. Por esta razão, se queremos obter uma resposta por parte do sistema de informação que nos permita satisfazer nossa necessidade informativa, há de

transformar essa formulação de maneira que o sistema possa entendê-la (PÉREZ GUTIÉRREZ, 2000, p. 247, tradução nossa)<sup>24</sup>.

Mas, o alcance da qualidade na informação recuperada também requer do planejamento de estratégias de busca específicas para cada tipo de base de dados, ou seja, “efetivamente a busca de informação, requer o desenvolvimento de uma efetiva e lógica estratégia de busca”<sup>25</sup> (DEBOWSKI, 2001, p. 372, tradução nossa).

Assim, para o planejamento da estratégia de busca, a identificação apropriada dos elementos descritivos e dos elementos temáticos de um item ou registro de informação contido em uma base de dados é de fundamental importância, ou seja,

Em princípio, a forma exata de uma estratégia de busca é determinada pela natureza da base de dados a ser acessada e pela sua respectiva estrutura de informação, isto é, pela formatação de seus campos de identificação do documento e dos campos de identificação do conteúdo temático do mesmo. A identificação dessa estrutura de campos de busca implica o conhecimento da documentação básica fornecida pelos produtores das bases e pelos bancos de dados onde estão hospedadas (LOPES, 2002a, p. 46).

---

<sup>24</sup> Normalmente, materializamos y representamos estos estados mentales mediante el enunciado de un language natual (castellano, catalán, inglés etc.). desgraciadamente, los SI no acostumbran a entender, por así decirlo, esas peticiones o consultas de información realizadas a partir de los enunciados de un language natural. Por esta razón, si queremos obtener una respuesta por parte del SI que nos permita satisfacer nuestra necesidad informativa, hemos de transformar esa formulación de manera que el sistema pueda entenderla.

<sup>25</sup> Effective information search requires the development of an effective and logical search strategy.

Uma estratégia de busca pode ser definida como a possibilidade parcial e o plano implícito que o pesquisador adota para resolver um problema de informação (TASSO et al, 2002). Todavia Lopes (2002b, p.61) afirma que “no âmbito da recuperação da informação, a estratégia de busca pode ser definida como uma técnica ou conjunto de regras para tornar possível o encontro entre uma pergunta formulada e a informação armazenada em uma base de dados”.

Neste sentido, é por isso que torna-se necessário a tradução do enunciado da linguagem natural, ou seja, do discurso, a uma linguagem determinada acessível pelo sistema de informação. A linguagem em questão se denomina **linguagem de interrogação**, e o resultado que obtemos com esta linguagem mediante a tradução recebe o nome de **equação de busca** (PÉREZ GUTIÉRREZ, 2000).

Desta forma, Péres Gutiérrez (2000), propõe uma gramática formal para a recuperação da informação. A princípio, uma gramática de uma linguagem concreta nos deve dar um conhecimento das propriedades sintáticas (relações que se produzem entre os signos dessa linguagem) e semânticas (relações que mantêm esse signos com os objetos que representam) que a determinam. Assim sendo, da mesma forma que as outras linguagens, a linguagem de interrogação que os sistemas de informação utilizam se encontra formada por um conjunto de símbolos que pode ser identificado como o léxico<sup>26</sup> da linguagem de interrogação. Este conjunto é normalmente constituído, por sua vez, por um conjunto de termos, pelos operadores booleanos e, ocasionalmente, por sinais de pontuação.

Pérez Gutiérrez (2000), assim, afirma que,

---

<sup>26</sup> O vocabulário, ou seja, o conjunto das palavras de uma língua (DICIONÁRIO AURÉLIO ELETRÔNICO, 2002).



- A **principal propriedade sintática** que possuem estas sucessões de símbolos é justamente a de ser uma equação de busca (se gramatical e estar corretamente constituída) e diferenciar-se, desta maneira, daquelas sucessões que não são. Ou seja, todas as equações de busca são sucessões gramaticais de símbolos do alfabeto da linguagem de interrogação, mas nem todas as sucessões de símbolos desse alfabeto são equações de busca e, portanto, gramaticais.
- Quando um usuário da linguagem de interrogação, ao propor uma equação de busca ao sistema de informação, o que pretende é expressar ou definir por compreensão certo conjunto de documentos que quer que o sistema de informação recupere e lhe permita consultar para satisfazer uma necessidade de informação original. Neste sentido, a **principal propriedade semântica** que possuem as equações de busca é a de expressar ou definir um conjunto de documentos.

Muitos sistemas de informação utilizam-se das lógicas de busca para se fazer a conversa (interação) entre pesquisador(es) e base de dados. A lógica de buscas constitui o meio de especificar combinações de termos que devam ser obtidas para se chegar a uma recuperação bem-sucedida.

A lógica booleana é resultante da aplicação da álgebra de Boole<sup>27</sup>. Segundo Pérez Gutiérrez (2000), este modelo é o que mais se destaca, perante os outros modelos dentro da literatura especializada.

---

<sup>27</sup> Álgebra binária (cujas variáveis podem assumir apenas dois valores), desenvolvida no século XIX pelo matemático inglês George Boole (1815-1864), que permite determinar se proposições lógicas são falsas ou verdadeiras, e que tem, atualmente, emprego fundamental em computação (DICIONÁRIO AURÉLIO ELETRÔNICO, 2002).

No modelo Booleano ou lógico, os termos da consulta são ligados por operadores lógicos [E], [OU], [NÃO] (AND, OR, NOT), sendo que neste modelo o sistema de recuperação da informação realiza a busca comparando seqüências de caracteres. Este modelo (assim como outros) também permite, eventualmente, contar com mecanismos com recursos de busca conforme apresentados por (RIBEIRO JÚNIOR, 2003a; ROWLEY, 2002; LARGE; MOUKDAD, 2001):

- **Truncamento**, que permite encontrar termos de raízes similares, ao empregar o caractere indicativo de truncamento, no começo ou no fim de uma palavra (por exemplo, a raiz bibliotec\* para encontrar as palavras Biblioteconomia; bibliotecário, biblioteca, bibliotecas);
- **Proximidade** ou **adjacência**, pois, muitas vezes a descrição de um assunto fica melhor com o emprego de uma expressão de duas ou mais palavras. É utilizado para encontrar dois ou mais termos em seqüência (por exemplo, para encontrar “Informação Científica” digita-se, Informação [ADJ] Científica);
- **Aninhamento**, o uso de parênteses (aspas ou outro tipo de pontuação), permite que os termos e operadores incluídos dentro dos parênteses, sejam tratados como um único termo – por exemplo, (Política [E] Economia) [NÃO] “Século XIX”. Tudo depende dos recursos que o mecanismo de busca da base de dados pode oferecer.

A seguir são apresentados os recursos de busca mais usuais.

<b>RECURSO</b>	<b>FUNÇÃO</b>
<b>Recursos de configuração, p. ex., ajuda, notícias, saída do sistema</b>	Configura o ambiente.
<b>Seleção dos termos de busca</b>	Permite a quem faz a busca identificar possíveis termos de busca ao examinar listas de termos de indexação ou Tesouros.
<b>Entrada de termos de busca</b>	Permite a quem faz a busca inserir os termos de busca.
<b>Combinação dos termos de busca</b>	Permite o desenvolvimento de estratégias de busca com o emprego da lógica de buscas.
<b>Especificação de campos</b>	Permite escolher os campos nos quais aparecem termos de busca (p. ex., título, resumo, texto completo, etc).
<b>Truncamento</b>	Permite que as buscas sejam feitas em cadeias de caracteres que não formam palavras completas (p. ex., Livr* para Livro, Livraria, Livreiro, Livros).
<b>Buscas por proximidade, adjacência e contexto</b>	Requer que as palavras estejam presentes em determinado contexto.
<b>Buscas e limitação de amplitude</b>	Realiza buscas numa faixa de amplitude (p. ex., data de publicação).
<b>Exibição dos resultados das buscas</b>	Mostra ao usuário quantos documentos e referências com os termos de busca foram encontrados.
<b>Exibição dos registros</b>	Mostra os registros na tela.
<b>Gerenciamento das buscas</b>	Permite rever as buscas feitas.
<b>Opções avançadas de exibição</b>	Para comportar registros em bases de dados de texto integral.
<b>Consultas conjugadas ou combinadas</b>	Agrega os resultados de buscas de grandes conjuntos de acervos dispersos de documentos.
<b>Exibição do tesauro</b>	Exibe os termos controlados de indexação e as relações entre os termos.
<b>Hipervínculos</b>	Permite aos usuários navegar entre ocorrências de termos em diferentes registros ou entre conceitos interligados por vínculos de hipermídia.

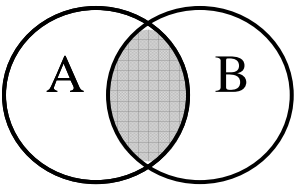
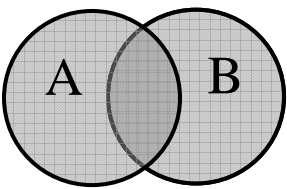
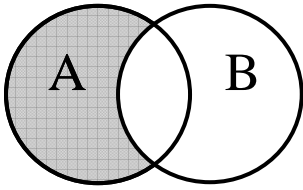
**Figura 11:** Recursos de busca usuais

**Fonte:** Rowley, 2002

Segundo Rowley (2002, p. 172), “os operadores Booleanos [...] também estão sujeitos a algumas variações”, por exemplo “**E**” pode ser representado em um mecanismo de busca de base de dados como “+”.

O modelo Booleano, é facilmente implementável e exato e, se o usuário possuir treinamento e souber o que quer, a expressividade da consulta é completa (RIBEIRO JÚNIOR, 2003a).

Veja o seu funcionamento:

Operador	Tipo de busca	Diagramas de Venn	Significado
<b>E</b>	Conjuntiva		Produto lógico, simbolizado por <b>A [E] B</b> . Ambos os termos <b>A</b> e <b>B</b> devem ser atribuídos ao documento para ele ser recuperado, p.ex., Imigração <b>[E]</b> Brasil <b>[E]</b> Século XIX, implica que todos esses termos foram atribuídos ao documento para que fosse recuperado.
<b>OU</b>	Aditiva		Soma lógica, simbolizada por <b>A [OU] B</b> . Um ou outro dos termos de indexação, <b>A</b> ou <b>B</b> , deve estar associado ao documento para ser recuperado. Este operador é comumente incluído quando <b>A</b> e <b>B</b> são considerados equivalentes para o objetivo da busca, p. ex., Cachaça <b>[OU]</b> Aguardente recuperaria os documentos a que houvesse sido atribuído o termo “Cachaça” ou Aguardente.
<b>NÃO</b>	Subtrativa		Diferença lógica, simbolizada por <b>A [NÃO] B</b> . O termo de indexação <b>A</b> deve ser atribuído, e atribuído na ausência do termo <b>B</b> , para que o documento seja recuperado, p. ex., Imigração <b>[NÃO]</b> Italianos pede todos os documentos sobre Imigração exceto aqueles onde “Italianos” também foi atribuído.

**Figura 12:** Operadores lógicos booleanos

**Fonte:** Rowley, 2002

Para uma equação de busca estar corretamente constituída, a sucessão dos termos de busca deve seguir o funcionamento acima. Devemos lembrar que os operadores **[E]** e **[OU]** são comutativos, ou seja,  $A [E] B = B [E] A$ ;  $A [OU] B = B [OU] A$ , já o operador **[NÃO]** não é pois,  $A [NÃO] B \neq B [NÃO] A$  (PÉREZ GUTIÉRREZ, 2000).

Pérez Gutiérrez (2000) ainda lembra que, uma sucessão de símbolos do alfabeto da Linguagem de Interrogação é uma equação de busca dessa linguagem se, e somente se, encontra gerada por um número finito de aplicações das seguintes **regras [sintáticas]**:

- (a) Se **T** é uma palavra ou uma frase da língua utilizada nos documentos, então **T** é uma equação de busca da Linguagem de Interrogação;
- (b) Se **A** e **B** são equações de busca da Linguagem de Interrogação então:
  - (b.I) **A [AND] B** é também uma equação de busca da Linguagem de Interrogação;
  - (b.II) **A [OR] B** é também uma equação de busca da Linguagem de Interrogação;
  - (b.III) **A [NOT] B** é também uma equação de busca da Linguagem de Interrogação.

A definição acima, nos gera o conjunto de todas as possíveis sucessões de símbolos do “alfabeto” da Linguagem de Interrogação que possuem a **propriedade sintática** de ser uma equação de busca dessa mesma linguagem.

E, se depois do processo de indexação de um documento **d**, obtém-se que **T<sup>1</sup>**, **T<sup>2</sup>**, **T<sup>3</sup>** são seus termos representativos (também denominados termos de indexação), pode-se identificar **d** da seguinte maneira:  $d = \{T^1, T^2, T^3\}$  (PÉREZ GUTIÉRREZ, 2000). Assim,

de acordo com Pérez Gutiérrez (2000) a definição, por compreensão, de um conjunto de documentos expressado por uma equação de busca da Linguagem de Interrogação que possui a **propriedade semântica**, se obtém a partir de um número finito de aplicações das seguintes **regras semânticas**:

(a) Se **T** é uma palavra ou uma frase da língua utilizada nos documentos gerenciados pelo Sistema de Informação, então:  $[T] = \{d: T \in d\}$ , onde o conjunto de documentos expressado pela equação de busca **T** se encontra formado por todos aqueles documentos **d** que incluem o termo **T** entre os termos que o representam.

(b) Se **A** e **B** são equações de busca da Linguagem de Interrogação, então:

(b.I)  $A [AND] B = \{d: d \in A \text{ e } d \in B\}$ , onde o conjunto de documentos expressado pela equação de busca **A [AND] B** está constituído por aqueles documentos **d** que pertencem ao conjunto representado pela equação **A** e que pertencem ao mesmo tempo ao conjunto expressado pela equação **B**.

(b.II)  $A [OR] B = \{d: d \in A \text{ ou } d \in B\}$ , no qual o conjunto representado pela equação de busca **A [OR] B** formam aqueles documentos **d** que pertencem ao conjunto expressado pela equação **A** ou ao conjunto expressado pela equação **B**. sendo que **[OR]** não está em um sentido excludente, ou seja, no conjunto definido por esta equação se encontram os documentos que pertencem a **A**, os que pertencem a **B** e os que pertencem simultaneamente a **A** e a **B**.

(b.III)  $A [\text{NOT}] B = \{d: d \in A \text{ e } d \notin B\}$ , onde mostra que o conjunto expressado pela equação de busca  $A [\text{NOT}] B$  se encontra constituído por aqueles documentos  $d$  que pertencem ao conjunto expressado pela equação  $A$  e que, por outro lado, não pertencem ao conjunto expressado pela equação  $B$ .

Destarte, vemos então que a cada uma das **regras sintáticas** temos uma **regra semântica** correspondente. Em definitivo, aproveita-se das regras sintáticas de formação das equações de busca para embasar ou fundamentar sobre esta definição da propriedade semântica e, desta maneira, tenta-se “[...] mostrar como o conjunto de documentos que expressa ou define, por compreensão, toda a equação de busca depende, sistematicamente, do conteúdo semântico das unidades significativas menores que a constituem”<sup>28</sup> (PÉREZ GUTIÉRREZ, 2000, p. 255).

Portanto, vemos que o trabalho de busca e recuperação da informação é um processo de diagnóstico, elaboração e fornecimento de resposta a um problema de informação de um usuário/cliente, através da tomada de decisões. É uma atividade altamente cognitiva, que exige o conhecimento e entendimento da questão ou situação-problema, a seleção e implementação de uma estratégia de busca e a obtenção da informação que responde à questão, no tempo mais hábil possível (MARTUCCI, 2000).

Um trabalho muito árduo que exige a constante adequação do especialista em busca e recuperação da informação (Bibliotecário) – visto que este é capacitado para entender

---

<sup>28</sup> [...] mostrar cómo el conjunto de documentos que expresa o define por comprensión toda ecuación de búsqueda depende sistemáticamente del contenido semántico de las unidades significativas menores que la constituyen.

a estrutura das bases de dados, a função dos campos e todos os demais mecanismos de busca, as estruturas de indexação e as estratégias de busca (DEBOWSKI, 2001). A seguir, no capítulo: Serviço de referência e a recuperação da informação pelo usuário final serão discutidos o papel do serviço de referência e as interações entre bibliotecário e usuário final.



## CAPÍTULO 4

A existência da linguagem é soberana, pois que as palavras receberam a tarefa e o poder de representar o pensamento.

FOUCAULT<sup>29</sup>

### SERVIÇO DE REFERÊNCIA E A RECUPERAÇÃO DA INFORMAÇÃO PELO USUÁRIO FINAL

O compartilhamento e uso coletivo da informação para o benefício da sociedade e dos indivíduos que a constituem, é a essência do conceito de biblioteca ou de qualquer outro tipo de unidade de informação. Foi para este compartilhamento ocorrer com maior eficácia que surgiu a partir de meados do século XIX, o **serviço de referência**, pois, devido ao enorme incremento da indústria editorial, da ampliação do ensino público e do avanço da alfabetização criou-se todo um novo público leitor. Segundo Grogan (1995, p.24),

Os estudos especializados expandiram-se, com os livros sendo publicados sobre assuntos cada vez mais específicos, e as bibliotecas aumentaram de tamanho. Chegara a hora em que os estudiosos já não mais conheciam nem mesmo os nomes de quem escrevia sobre os temas de seu interesse, e então começaram a procurar os livros pelos seus assuntos.

É na assistência a este novo e exigente público leitor, numa nova estrutura de biblioteca (mantida por impostos), é que se podem localizar as origens do que podemos chamar de serviço de referência (GROGAN, 1995). E hoje,

---

<sup>29</sup> FOUCAULT, Michel. **As palavras e as coisas**. São Paulo: Martins Fontes, 1999. 541 p.

Os serviços de referência possuem um lugar central na biblioteca e nos serviços de informação. São também considerados como serviços personalizados já que, na maioria dos casos, uma discussão pessoal acontece entre um usuário e um bibliotecário de referência (CHOWDHURY, 2002, p. 258, tradução nossa).<sup>30</sup>

Através dos serviços de referência, o bibliotecário não só tenta entender a informação específica que o usuário precisa, mas também coleta informações sobre o usuário, como, por exemplo, o conhecimento do assunto procurado pelo usuário, o propósito da informação específica encontrada, e assim por diante. Desse modo, o bibliotecário de referência pode filtrar, muitas vezes, a informação recuperada selecionando as fontes mais apropriadas para um determinado usuário em um determinado ponto do tempo. (CHOWDHURY, 2002).

Assim sendo, os usuários das bibliotecas, auxiliados pelo bibliotecário de referência têm melhores condições, de modo mais perfeito e completo, de aproveitarem o acervo de uma biblioteca do que o fariam sem essa assistência (GROGAN, 1995). Ou seja, o serviço de referência foi desenvolvido para que todos os recursos informacionais da unidade de informação fossem aproveitados da melhor forma possível. É neste setor, que também se dão as atividades e os processos de recuperação da informação.

Michel Menou e Claire Guinchat (1994, p.18), em livro básico para a área da Biblioteconomia e Ciência da Informação, dizem, com muita pertinência, que

---

<sup>30</sup> Reference services have taken a central place in library and information services. They are also regarded as personalised services since in most cases a personal discussion takes place between a user and a reference librarian.

Os objetivos das atividades documentais [e biblioteconômicas] são selecionar na massa de informações veiculadas, os elementos de conhecimento, fornecer a qualquer pessoa as informações de que ela necessita, no momento que as solicita, e ainda conservar estas informações atualizadas, sem alterá-las.

De acordo com estes autores, os objetivos de todas as atividades biblioteconômicas consistem em recuperar a informação necessária para a resolução dos problemas apresentados, ou melhor, diante de uma problemática qualquer, há a necessidade de se obter os elementos que diminuam as incertezas e os conflitos, fornecendo-os aos usuários e/ou clientes, da melhor forma possível, ou seja, agregando valor à informação, para que assim os objetivos pretendidos destes últimos possam ser alcançados.

Mas, foi partir do momento em que o foco das atividades da Biblioteconomia deixou de ser o acervo/documento e passou a ser a informação, que ocorreu a grande mudança de paradigma em nossa área. Mudar um paradigma envolve mudar os pressupostos do conhecimento que o determina e também as crenças, os valores e a ideologia a ele, associados (LEITÃO; ROSSI, 2000).

As mudanças estão afetando de maneira mais complexa os modelos tradicionais de trabalho dos bibliotecários, porque o objeto de trabalho destes é a informação. A informação, portanto, com o surgimento de uma preocupação quanto ao processamento de informações necessárias à tomada de decisões em qualquer ambiente, tem sido afetada sistematicamente pelas tecnologias de informação e comunicação, modificando seu formato, seu suporte, seu processamento, recuperação e disseminação e influenciando na forma de mediação entre o profissional da informação e o usuário/cliente (VALENTIM, 2000; COHEN, 2002).

Segundo Bains (1997), o rápido desenvolvimento em tecnologia, teve um grande efeito na comunicação, armazenamento e recuperação da informação que, no final do século

XX, tem sido chamado de ‘sociedade da informação’. A informação está disponível em muitos formatos e em muitas variadas fontes do que meio século atrás e, por conseguinte, a revolução eletrônica teve um massivo efeito em nossa habilidade para armazenar e recuperar a informação.

Com o desenvolvimento das tecnologias de informação e comunicação, as bibliotecas passaram por grandes mudanças em seu ambiente organizacional. A coleção deixou de ser uma propriedade individual de cada biblioteca, passando a ser uma propriedade coletiva, aumentando a necessidade de programas cooperativos (FIGUEIREDO, 1999). Correa da Cruz et al. (2003, p. 52), neste sentido, ainda aponta que,

A nova realidade aponta também para o aumento do número de consórcios e para o fortalecimento da filosofia do “acesso ao invés da propriedade”, como forma de tornar a informação mais acessível, promovendo o avanço cultural e científico.

Do acervo constituído apenas por materiais de papel (principalmente livros), passou a oferecer também materiais em CD-ROM (bases de dados, anais, suplementos e até mesmo livros em CD-ROM), além do acesso à rede de computadores para dados compartilhados (acesso via rede) e o acesso à informação hipermídia disponível on-line na Web (qualquer lugar a qualquer tempo). Atualmente, “a idéia não é mais adquirir o que o usuário pode querer, mas obter acesso a esse material, independentemente de seu formato e do lugar em que se encontre” (FIGUEIREDO, 1999, p. 12).

Pode-se perceber que desde o surgimento dos computadores, bibliotecas e bibliotecários têm utilizado os computadores para serviços básicos como a catalogação, a indexação e a organização do acervo. Mas, foi o acesso **on line** aos bancos de dados por meio das redes de telecomunicações, que permitiu a dinamização dos processos de recuperação e

disseminação da informação (OHIRA; PRADO, 2002). Ou melhor, a tecnologia é um catalisador de mudanças particularmente importantes e urgentes para as bibliotecas, uma vez que cria novas necessidades e altera velhos e sólidos paradigmas estabelecidos ao longo de muitos séculos (LEVACOV, 1997).

Conforme ressaltam Modesto e Macedo (1999, p. 64),

[...] a tecnologia e os processos inovadores ainda estão transformando o ambiente da informação, indicando que reavaliações e re-adequações deverão ser continuamente feitas do convencional às novas e futuras realidades.

Além disso, com a queda dos custos de produção e dos avanços das tecnologias de informação e comunicação, a utilização de bases e bancos de dados on-line por usuários finais vem crescendo a cada dia, de forma que, o acesso, a busca e a recuperação da informação têm-se deslocado do terreno dos especialistas da informação (bibliotecários) para o dos usuários finais (CARO-CASTRO; TRAVIESO RODRÍGUEZ; CEDEIRA SERANTES, 2003). Isso também tem um profundo efeito no papel do intermediário bibliotecário, uma vez que, muitas pessoas têm se tornado familiares com o crescente do número de bases de dados amigáveis para usuários (BAINS, 1997).

Segundo Mercado (1999, p. 259, tradução nossa), “Há vinte e cinco anos atrás [1974], bibliotecários conduziam buscas em bases de dados on line para os usuários de bibliotecas, normalmente a um custo por recuperação”<sup>31</sup>. Porém, o processo de “desintermediação” tem estado em operação desde os anos sessenta, quando o usuário-final pesquisando nos primeiros bancos de dados on-line se tornou um fenômeno significativo.

---

<sup>31</sup> Twenty-five years ago librarians conducted online database searches for library users, usually on a cost-recovery basis.

Segundo Sturges (2001, p. 62, tradução nossa),

A razão por trás disto era simples. Como as informações ficaram disponíveis em terminais de computador unidos por modem a bancos de dados remotos, a importância da biblioteca como um lugar onde a informação era armazenada foi diminuída.<sup>32</sup>

Ou seja, o difundido alcance e uso da Internet tem apresentado um dobrado desafio para o familiar paradigma da referência: a elevação dos recursos, tanto livres como licenciados, disponíveis pela Internet, e o uso de numerosas tecnologias baseadas na Internet com possíveis modos de interagir com pessoas que apresentam questões e necessidades de informação (JANES, 2002).

Assim, a barreira (principalmente a de custo por recuperação) que uma vez fez os produtos e serviços de informação on-line um domínio do bibliotecário não existe mais (MERCADO, 1999). Embora o bibliotecário seja o profissional quem tem como uma de suas especialidades a busca e a recuperação da informação e seja, a princípio, o mais habilitado para a realização desta atividade, há uma forte tendência de que o usuário tenha, cada vez mais, independência na identificação e acesso a informação de que precisa. Além disso, sendo o bibliotecário o principal intermediário entre o usuário e informação, haveria uma demanda muito alta a ser suprida. Pois, a explosão dos números e tipos de recursos de informação eletrônica, confronta este profissional com uma miríade de recursos online, em CD-ROM, em base de dados de texto completo, e em outros recursos de informação digital, como os repositórios institucionais e as bibliotecas digitais (MA, 2002).

---

<sup>32</sup> The rationale behind this was simple. As information became available at computer terminals linked by modems to remote databases, the importance of the library as a place where information was stored was diminished.

Isto é, desde que a Web e as bibliotecas digitais<sup>33</sup> tornaram-se um fenômeno significativo, provendo acesso direto a fontes de informação e serviços sem a intervenção de intermediários humanos (principalmente os bibliotecários), uma pergunta pertinente a esse respeito é se os usuários/clientes ainda precisam de serviços de referência em bibliotecas digitais e, neste caso, como melhor oferecer tais serviços (CHOWDHURY, 2002).

Nota-se, portanto, uma outra mudança de paradigma,

[...] do acesso à informação através do modelo centrado na informação, para o modelo centrado no usuário. Isso quer dizer que o acesso à informação, não é dirigido pela estrutura do sistema (uma biblioteca, uma base de dados) mas sim pelas visões da biblioteca e das bases de dados necessárias para satisfazer uma necessidade da informação, como percebida pelo usuário (FIGUEIREDO, 1999, p. 13).

Esta mudança de foco do acesso a informação, dos modelos centrados na informação para os centrados no usuário, parte do princípio de que a necessidade de informação de um usuário é específica àquele indivíduo e que “uma necessidade de informação não pode ser separada da situação que a criou e do indivíduo que a percebeu” (FIGUEIREDO, 1999, p. 13).

---

<sup>33</sup> Biblioteca digital – [...] é um serviço de informação no qual todos os recursos informacionais estão disponíveis em formato processável por computador, ou seja, o armazenamento, preservação, recuperação, acesso e apresentação das informações ocorrem através do uso de tecnologia digital (discos ópticos e magnéticos). Neste sentido, a biblioteca digital não contempla materiais convencionais impressos como livros, já que estes seriam convertidos/digitalizados para o formato digital. A informação é pois compartilhada simultânea e instantaneamente por meio de acesso local ou remoto, já que a biblioteca digital se estrutura em redes de computadores, que são também veículos digitais. Este é o ponto chave da biblioteca digital: sua informação pode ser acessada remotamente em múltiplas vias (MODESTO; MACEDO, 1999, p. 64).

Observa-se aqui um problema para o usuário, o qual, seu sucesso no acesso à informação depende de vários fatores, incluindo quão bem entende a linguagem, a estrutura dos dados na base, e quanto bem entende qual(is) dos dados disponíveis satisfazem sua necessidade de informação (FIGUEIREDO, 1999).

Deste modo vemos que possibilidade de acesso não é garantia de sucesso na obtenção da informação necessária, pois “A utilização inadequada de uma nova tecnologia, como consequência de uma reduzida compreensão dos recursos que a mesma pode proporcionar, contribui para que ela não seja utilizada de forma ampla” (DIAS, 2002, p. 19). Isto é, “a simples existência de ferramentas não garante de maneira alguma a sua [adequada] utilização” (DIAS, 2002, p. 21).

E, Lopes (2002b, p. 65), diz ainda que,

Apesar dos intensivos programas de treinamento oferecidos pelos produtores das bases de dados, pelos próprios sistemas de recuperação em linha, de toda a documentação existente sobre as características de cada base de dados e suas respectivas estruturas de informação, dos sistemas amigáveis que oferecem “menus” para guiar o usuário em cada etapa do processo de busca, das linguagens de busca com recursos especiais para se aproximarem cada vez mais do usuário inexperiente, o processo de busca continua sendo um fator de dificuldade que ainda não foi minimizado pelas novas tecnologias disponíveis.

TASSO et al. (2002, p. 343, tradução nossa), salienta que

[...] hoje muitos sistemas de recuperação da informação são diretamente usados por usuários finais. Entretanto, por causa da ausência do



intermediário humano [bibliotecário], os usuários enfrentam uma tarefa difícil, porque eles têm que saber como interagir diretamente com o sistema e contender com vários tipos de problemas: aprender comandos de interface, usar a lógica booleana, escolher os termos mais efetivos e planejar e aplicar estratégias de busca efetivas.<sup>34</sup>

Assim sendo segundo Foster et al (2002, p. 884, tradução nossa)

Nas duas décadas passadas, uma segunda abordagem tem se desenvolvido na Recuperação da Informação: a de que os usuários e intermediários interagem com sistemas de Recuperação de Informação. Para usuários e intermediários, o problema de encontrar documentos úteis em um sistema consiste de formar um entendimento do problema do usuário; e traduzir aquele entendimento em uma questão para ser apresentada ao sistema de informação<sup>35</sup>.

Saracevic et al (1990) ainda sugere as categorias para a análise do discurso das interações usuário-intermediário como sendo:

---

<sup>34</sup> Today many information retrieval systems are directly used by end users. Because of the absence of a human intermediary, users face a difficult task, because they have to know how to interact directly with the system and cope with various kinds of problems: choosing the more effective terms, and designing and applying effective search strategies.

<sup>35</sup> In the past two decades, a second research approach has developed in IR: that of users and intermediaries interacting with IR systems. For users and intermediaries, the problem of finding useful documents from an IR system consist of forming an understanding of user's problem; and translating that understanding into a query to be presented to the information system.

- 1) **Contexto** – o problema ou tarefa do usuário à mão; estágio de busca da informação; informação.
- 2) **Terminologia** – elaboração de restrições e na modificação de conceitos, de termos, de palavras-chave e de descritores; geração de termos; especificação do assunto; restrições como com respeito a idioma, termos técnicos, soletração.
- 3) **Explicações dos sistemas** – funcionamentos e aspectos técnicos dos sistemas usados e explicações técnicas de busca; características das bases de dados e documentos no sistema.
- 4) **Táticas de busca** – seleção e variação de termos, campos, morfologia, lógica das buscas; comandos; seleção e variação.
- 5) **Revisão e relevância** – revisão da busca com respeito à resposta; avaliação das fontes ou conteúdos; julgamentos de relevância da realimentação (feedback) das respostas.
- 6) **Ação** – descrição de uma atividade contínua ou iminente.
- 7) **Estímulos e pausas** – As comunicações estimulam maiores satisfações, reconhecimentos, expressões de contentamento etc., indicando envolvimento dos ouvintes, for exemplo, “O.K”, “Nossa!”, “Huhum”, “Certo”.

**8) Estranhezas** - expressões vocais estranhas à interação de busca.

E Saracevic (1990) também sugere as categorias para a análise de busca nas interações computador-bibliotecário:

- 1) **Arquivo** – seleção da base de dados;
- 2) **Terminologia** – seleção dos termos;
- 3) **Restrição** – linguagem; limitações de ano;
- 4) **Explicação e revisão** – revisão de comando, resultados;
- 5) **Respostas** – exibição dos resultados, escolha do formato ou extensão;
- 6) **Inativo** - nenhuma atividade do computador.

Deve-se ter em mente que a interação na recuperação da informação consiste de uma série de interações e adaptações dinâmicas entre níveis. Como a interação progride, mudam-se as coisas. Por exemplo no nível superficial pode ser mudada uma questão, termos podem ser adicionados ou retirados, diferentes táticas podem ser empregadas, transmitindo e afetando mudanças para os outros níveis (FOSTER, 2002). Lembrando-se que a interação pode ser também entendida como aquela entre os diferentes estratos do usuário e dos computadores ou os níveis realizados na superfície da interface de busca. Do lado do usuário, pode ser o modelo de níveis o superficial, o cognitivo, o afetivo e o situacional. No lado do “computador” pode ser o modelo de níveis: o superficial, o Mecanismo de acesso, o processamento, e o conteúdo (SPINK; SARACEVIC, 1997).

A presença do bibliotecário nos processos de busca e recuperação da informação torna-se importante visto que, ao longo de muitos anos, nenhum treinamento foi

providenciado para os usuários finais, enquanto que os bibliotecários foram aprendendo as complicadas linguagens de comandos e estratégias de busca (MERCADO, 1999).

Os dois, usuário e intermediário (bibliotecário), possuem conhecimentos e habilidades diferentes, mas importantes para o processo de busca :

[...] o intermediário que operacionaliza as buscas dispõe de conhecimentos prévios sobre o uso da linguagem controlada e da linguagem natural, porém o usuário final possui uma profunda intuição sobre a terminologia de seu campo de especialização [...] (LOPES, 2002a, p. 47).

Nos processos de recuperação da informação, o bibliotecário ainda é o profissional que dispõe de maior habilidade e conhecimento para utilizar tanto a linguagem controlada como a natural, sendo capaz de traduzir a necessidade de informação do cliente/usuário tanto para a linguagem de busca como para as características de cada base (LOPES, 2002a; 2002b). É ainda o melhor profissional capacitado na intermediação, busca, acesso e análise de informações, ou seja, que tem a competência para

[...] responder/suprir demandas de informação, prover serviços de informação, realizar pesquisas e consultoria em informação, fazer buscas on-line, acessar/obter documentos, transferir, gerenciar ou comercializar informações, coletar informações para terceiros (MARCHIORI, 1998, p. 4).

Logo, pensar que os agentes<sup>36</sup> ditos “inteligentes” irão substituir os intermediários humanos da informação, no caso, os bibliotecários, é bastante improvável já que ambos têm

---

<sup>36</sup> Ver **Capítulo 2**, sub-item: 2.3 Recuperação da Informação e as Tecnologias da Informação e Comunicação.

diferentes qualidades e capacidades. É mais provável a sua íntima cooperação e uma alteração de tarefas que ambos os tipos executam. Os sistemas de agentes inteligentes intermediários ou sistemas especialistas, irão executar tarefas padrões e todas essas tarefas que um programa de computador pode fazer mais rapidamente ou até melhor do que um humano. Já os intermediários humanos irão tratar dos problemas mais complicados, pois, “a Internet continua a crescer, e certamente, continuará a crescer. A grande ameaça que isto acarreta é que a Internet se irá tornar muito grande e diversa para que um humano possa, sozinho, nela trabalhar corretamente” (NOGUEIRA, 2000, p. 5). Cabe ao bibliotecário dominar e utilizar-se de tais ferramentas, pois, “esses sistemas têm aplicação na recuperação da informação, na indexação, no resumo, na classificação, na catalogação e na referência” (FIGUEIREDO, 1999, p. 16).

Além disso, como destaca Morales (1995), dois trabalhos que tem que se destacar, por um lado, o papel do bibliotecário ao realizar análise documentária para detectar a relevância dos documentos (impressos ou eletrônicos) assim como sua habilidade para selecionar e implementar **interfaces satisfatórias** para o usuário a quem serve.

Desta maneira, o bibliotecário deverá também “[...] trabalhar conjuntamente com outros colaboradores da área da informática na criação de softwares referenciais, ou seja, os agentes inteligentes” (MODESTO; MACEDO, 1999, p. 68).

Ou melhor, o domínio e conhecimento das ferramentas tecnológicas, usadas para promover a difusão da informação são, portanto, considerados estratégicos e críticos, colocando aquele bibliotecário que detém estas noções em situação privilegiada no futuro (JAMIL; NEVES, 2000).

Vê-se, então, que o papel do bibliotecário, especialista da informação na referência, não é somente o de oferecer respostas aos clientes/usuários, mas também prepará-los para resolver efetivamente suas necessidades informacionais, propiciando-os a formar um

pensamento crítico sobre suas fontes de pesquisa (KASOWITZ, 1998 citado por ARELLANO, 2001, p. 9). Para Modesto e Macedo (1999, p. 62),

Ao bibliotecário, cabe manter-se continuamente atualizado sobre as tecnologias que afetam seu ambiente de trabalho. Enquanto mediador, entre o usuário e a informação, ele tem a responsabilidade não apenas de fornecer, mas orientar o usuário no uso das tecnologias que armazenem a informação.

Torna-se, portanto, necessário que o bibliotecário, na mediação entre o usuário e a informação desejada, exerça a assessoria, gerenciando e dando informações, além de praticar formas de treinamento na utilização de bases e bancos de dados com qualidade, valor agregado, precisão e rapidez, de forma a estabelecer conexões entre aqueles que detêm a informação e os que a querem (MODESTO; MACEDO, 1999).

E também, segundo Cuenca (1999, p. 291),

Para que os usuários conheçam os sistemas automatizados de recuperação da informação, sejam capazes de elaborar uma estratégia de busca simples e saibam utilizar a “nova biblioteca eletrônica” e os recursos informacionais de que dispõem, é necessário que as bibliotecas ofereçam treinamentos e cursos específicos, como modalidades de programas educativos.

Conforme ressalta Debowski (2001), o papel do profissional no desenvolvimento das habilidades de busca e recuperação da informação dos usuários finais dos serviços de informação, tornou-se ainda mais necessário. Isto é, o bibliotecário tem papel fundamental nesse processo, uma vez que cabe a ele racionalizar, divulgar e treinar os usuários para melhor

uso desses importantes recursos (CORREA DA CRUZ et al, 2003). Há, neste sentido, a necessidade de se buscar novos métodos para melhorar os serviços de referência (JANTZ, 2002, p. 62). Entretanto, “em cada caso é possível repensar significados e na transformação da prática da referência, mas qualquer mudança deve ter lugar dentro do contexto da referência como atualmente é praticada, sentida e entendida”<sup>37</sup> (JANES, 2002, p. 549, tradução nossa).

#### **4.1 Abordagens na Recuperação da Informação**

A necessidade de avaliações e estudos sobre as buscas feitas pelos usuários-finais, deveria tornar-se parte integrante dos serviços de referência. Alguns pesquisadores já perceberam a relevância de se dedicar ao estudo do tema. A necessidade de recuperar informações promoveu problemas e pesquisas exploratórias de fenômenos, processos e variáveis, causas e efeitos, comportamentos e manifestações,

[...] levando a estudos do comportamento humano frente à informação; a interação homem-computador, com ênfase no lado humano da equação; relevância, utilidade, obsolescência e outros atributos do uso da informação juntamente com medidas e métodos de avaliação dos sistemas de recuperação da informação; economia, impacto e valor da informação, dentre outros (SARACEVIC, 1996, p.45).

Nos estudos sobre recuperação da informação há duas linhas principais de investigação. A primeira é focada nos sistemas de recuperação da informação algorítmicos, ou

---

<sup>37</sup> In each case, there is potential for significant rethinking and transformation of the reference enterprise, but any such changes must take place within the context of reference as currently practiced, envisioned and understood.

seja, utiliza-se uma abordagem computacional para pesquisas sobre recuperação da informação. Segundo Ribeiro Júnior (2003b),

A expressão abordagem computacional pode ser entendida como toda abordagem em pesquisa sobre Recuperação Informação que envolve o uso de um sistema de recuperação de informação baseado no uso de um software, que inclui uma interface de busca, um banco de dados (no sentido que é dado na área de computação) e um mecanismo de busca.

Neste enfoque as pesquisas, geralmente, são realizadas por pesquisadores e profissionais da área da Ciência da Computação. Mas pode-se verificar na literatura, autores da área da Ciência da Informação investigando este mesmo enfoque (RIBEIRO JÚNIOR, 2003b).

O segundo é aquele centrado no usuário, o qual se estuda o comportamento de busca dos usuários de sistemas de recuperação de informação, como descreve Dias (2003, p. 9),

Desde a década de 90, a preocupação com a qualidade da informação ofertada na Internet tem suscitado vários estudos na área da Ciência da Informação e de interação Homem-Computador, principalmente no estabelecimento de diretrizes e métodos de avaliação, como tentativas de garantir a confiabilidade da informação e proporcionar uma experiência eficiente e agradável ao usuário, independentemente tipo de portal visitado (busca de informações, comércio eletrônico, home-banking, governo eletrônico).



Na abordagem computacional, com o enfoque nos sistemas de recuperação da informação, é o mecanismo que permite armazenar e recuperar a informação, a estratégia de busca é a forma de entrada no sistema e os documentos recuperados são os resultados. Esta delimitação contrasta com a situação que acontece no mundo real, onde as necessidades de informação surgem em contextos concretos e as pessoas estão implicadas no processo de busca. A perspectiva centrada no usuário amplia os limites do Sistema de Recuperação da Informação com a incorporação do componente humano e do ambiente em que acontece a busca por informações. (CARO-CASTRO; TRAVIESO RODRÍGUEZ; CEDEIRA SERANTES, 2003).

Conforme discorre Robins (2000, p. 57, tradução nossa),

Desde meados do Século XX, a maioria dos esforços para melhorar a recuperação da informação tem focado em métodos que comparam representações de textos com representações de questões. Recentemente, porém, pesquisadores têm empreendido na tarefa de entender o papel do humano, ou usuário, na recuperação da informação. A suposição básica entre estes esforços é que nós não podemos desenhar sistemas de recuperação efetivos sem um pouco de conhecimento de como os usuários interagem com eles. Então esta linha de pesquisa que estuda os usuários no processo de consulta direta a um Sistema de Recuperação da Informação, é chamada Recuperação Interativa da Informação.<sup>38</sup>

---

<sup>38</sup> Since the middle of the 20th century, most efforts to improve information retrieval have focused on methods of matching text representations with query representations. Recently, however, researchers have undertaken the task of understanding the human, or user, role in IR. The basic assumption behind these efforts is that we cannot design effective IR systems without some knowledge of how users interact with them. Therefore, this line of research that studies users in the process of directly consulting an IR system is called *interactive information retrieval* (IIR).

Ou seja, os pesquisadores começaram a olhar para o outro lado da equação na Recuperação da Informação: seres humanos que usam sistemas de recuperação da informação (ROBINS, 2000).

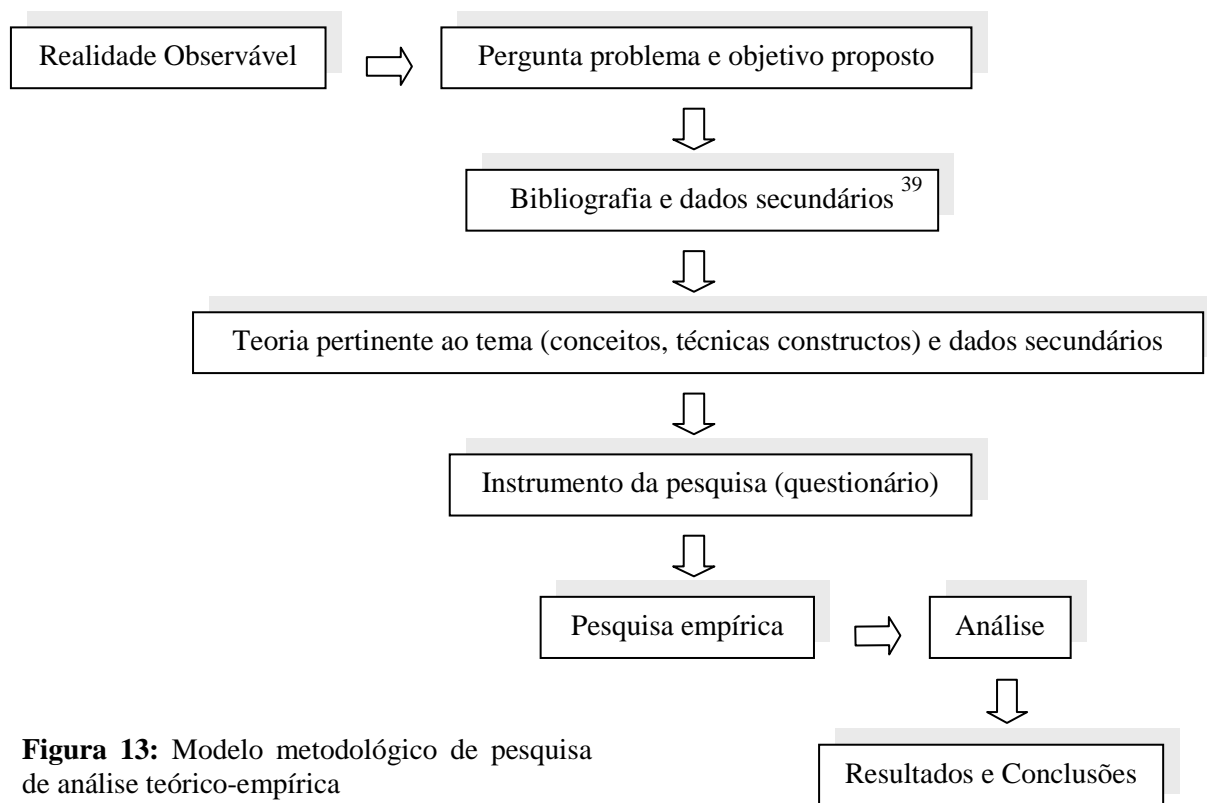
O próximo capítulo explicará os procedimentos metodológicos adotados para o desenvolvimento desta pesquisa.

## CAPÍTULO 5

O único modo de escapar da corrupção causada pelo sucesso  
é continuar a trabalhar.  
ALBERT EINSTEIN.

### PROCEDIMENTOS METODOLÓGICOS

Esta pesquisa é de caráter teórico-empírica e de nível exploratório, ou seja, de análise interpretativa de dados obtidos diretamente no campo ou origem dos eventos pesquisados (dados primários), com apoio bibliográfico, que possibilita uma abordagem qualitativa e quantitativa dos dados. A análise qualitativa caracteriza-se por permitir analisar as interações entre variáveis e possibilitar a compreensão de comportamentos e, a análise quantitativa caracteriza-se pelo emprego de procedimentos matemáticos e estatísticos para coleta e tratamento dos dados.



**Figura 13:** Modelo metodológico de pesquisa de análise teórico-empírica

**Fonte:** Mendes; Tachizawa, 2001

<sup>39</sup> “Dados secundários [...] são aqueles obtidos, por exemplo, de obras bibliográficas ou de relatórios de pesquisas anteriores sobre o tema” (MENDES; TACHIZAWA, 2001. p. 53).

Assim, para identificar as contribuições que o profissional bibliotecário pode proporcionar para a busca e recuperação da informação, na interação com o usuário, utilizou-se de **levantamento bibliográfico** de documentos primários, secundários e terciários nos idiomas espanhol, inglês e português para leitura, seleção, anotação dos conceitos aplicáveis ao tema e, conseqüentemente, a redação dos capítulos do relatório parcial e final. Compreende-se por fontes bibliográficas primárias, obras produzidas com a interferência direta do autor como, livros, teses, Trabalhos de Conclusão de Curso (TCCs), relatórios técnicos, normas técnicas, artigos científicos e outros. Como fontes secundárias, documentos produzidos a partir de documentos primários e são representadas por exemplo pelas enciclopédias, tratados, manuais, dicionários e outros. Como fontes terciárias documentos que tem por função guiar o usuário para as fontes primárias e secundárias como, bibliografias, catálogos coletivos, guias de literatura, diretórios, índices e outros (MUELLER, 2000).

O levantamento bibliográfico dos documentos primários, secundários e terciários, para a revisão bibliográfica, foi feito nas seguintes fontes bibliográficas da área de Ciência da Informação:

- Livros, periódicos, anais e relatórios;
- Bases de dados nacionais e internacionais, textuais e referenciais (Web of Science, WebSPIRS, Current Contents, Probe, Scielo, Cruesp, Emerald-Library, CAPES, dentre outras).
- Internet.

A partir dos procedimentos metodológicos adotados, pretende-se atingir os objetivos intermediários descritos abaixo;

- Identificar o perfil e as necessidades informacionais dos alunos da Pós-Graduação da UNESP – Marília,

- Verificar quais as bases de dados disponíveis para consulta são conhecidas e utilizadas eficazmente pelos alunos da Pós-Graduação da UNESP – Marília,
- Identificar as necessidades de otimização dos processos de planejamento e operacionalização das estratégias de busca para uma efetiva recuperação da informação pelos clientes/usuários;

Para isto, a coleta de dados foi desenvolvida em 3 etapas, sendo elas:

- a) Caracterização da Infra-estrutura para a pesquisa bibliográfica na Faculdade de Filosofia e Ciências (UNESP-FFC-Marília);
- b) Aplicação de questionário aos alunos sujeitos da pesquisa (alunos da pós-graduação da UNESP-FFC-Marília;
- c) Observação sistemática do processo de busca dos alunos sujeitos da pesquisa.

A primeira etapa (**a**), constituiu no levantamento da infra-estrutura disponível (bases disponíveis, equipamentos, treinamentos, pessoal disponível, serviços oferecidos pela biblioteca aos alunos-sujeitos da pesquisa), junto ao setor de referência da biblioteca da unidade, mediante o envio (por e-mail) de um questionário à bibliotecária responsável pelo setor de referência.

As etapas seguintes (**b** e **c**) dizem respeito à coleta de dados junto aos sujeitos da pesquisa e visavam identificar às necessidades informacionais, bem como o conhecimento e dificuldades na utilização das bases e dados disponíveis para consulta, bem como identificar as necessidades de otimização dos processos de planejamento e operacionalização da estratégia de busca para uma efetiva recuperação da informação pelos clientes/usuários.

Esta pesquisa tem como sujeitos os alunos da Pós-graduação da UNESP-FFC-Marília. Optou-se por esse segmento, uma vez que estes, para a realização de suas dissertações, teses, papers, etc, necessitam de levantamentos bibliográficos pertinentes aos seus temas de pesquisa.

A UNESP-FFC-Marília conta, atualmente, com quatro Programas de Pós-Graduação (dentro do campus) nas áreas de: **Ciência da Informação** (com 23 alunos); **Ciências Sociais** (com 58 alunos); **Educação** (com 145 alunos), sendo 78 do Mestrado e 67 do Doutorado e **Filosofia** (com 18 alunos), somando-se em um total de 244 alunos matriculados no ano de 2004 <sup>40</sup>, como pode-se verificar na **Tabela 1**.

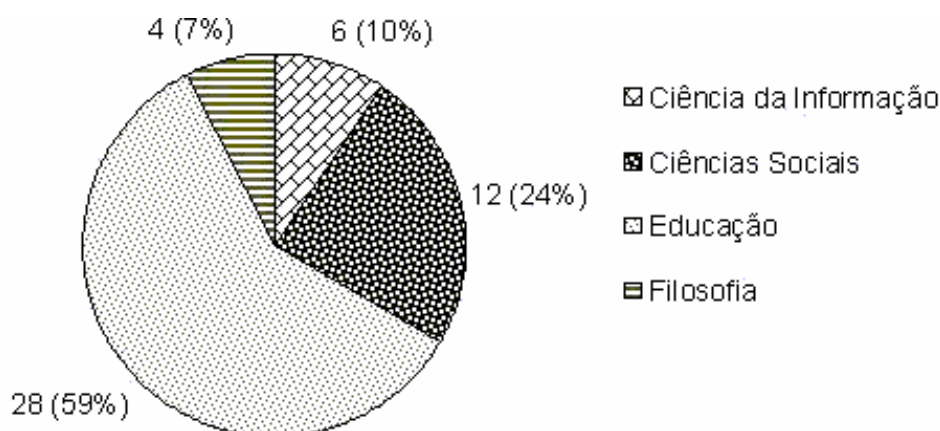
**Tabela 1 – Número de alunos matriculados nos programas de Pós-Graduação da Unesp Marília em 2004**

Programa	Número de alunos	%
Educação	145	59
Ciências Sociais	58	24
Ciência da Informação	23	10
Filosofia	18	7
Total	244	100

Estabeleceu-se então que uma amostra de 20% deste total, ou seja 50 sujeitos, participariam desta pesquisa, respondendo a um questionário (**etapa b**). Por se tratar de um grupo composto por quatro programas de Pós-Graduação com número de alunos matriculados bastante diferenciado, optou-se pela composição de uma amostra estratificada proporcional da população. Segundo Vieira (1999, p. 17), “usa-se o processo de amostragem estratificada quando a população se apresenta dividida em estratos, isto, é, quando a população está dividida em grupos distintos” como é o caso desta pesquisa.

<sup>40</sup> Dados obtidos junto à Seção de Pós-Graduação no início do ano letivo de 2004.

O número de 50 sujeitos (20% do total de alunos) foi dividido entre os quatro estratos da maneira como ilustra o **Gráfico 1**:



**Gráfico 1:** Número de sujeitos por amostragem estratificada proporcional para aplicação de questionário

O levantamento de dados referiu-se às necessidades informacionais, bem como o conhecimento e dificuldades na utilização das bases e dados disponíveis para consulta (**etapa b**). Utilizou-se da aplicação de **questionário**, com questões abertas e fechadas (apêndice 1) para os mesmos, no qual procurou-se extrair dados que pudessem ser analisados qualitativamente e quantitativamente pois, conforme Caro; Travieso e Cedeira (2003, p. 45, tradução nossa), nas pesquisas sobre recuperação da informação com a perspectiva centrada no usuário, “[...] se utilizam simultaneamente métodos qualitativos e quantitativos, o que demonstra a tendência da integração das duas orientações metodológicas”.

Vale ainda salientar que,

Uma análise de pesquisas anteriores de buscas de usuários finais, tem mostrado que os questionários, como um método de coleta de dados tem provado ser consistentemente conhecido. [...] Oferecem dados subjetivos, que podem ser usados para avaliar o que as pessoas percebem em cada caso.

Isto pode ser útil se a percepção do usuário for uma parte integrante da pesquisa (BAINS, 1997, p.4).

A aplicação de um teste piloto, foi feito com quatro alunos da Pós-graduação da UNESP-FFC-Marília. Após as devidas análise e adequações o questionário foi aplicado a referida amostra de 20 % da população total (50 sujeitos), por meio de um sorteio, respeitando a proporcionalidade estabelecida.

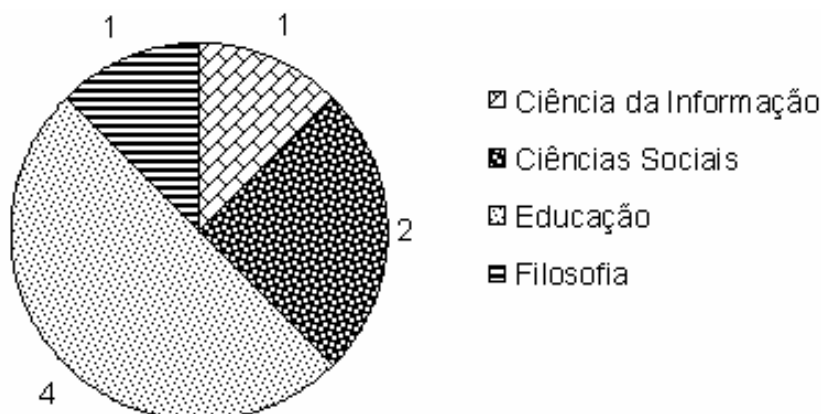
Decidiu-se pela aplicação de questionário auto-aplicado via e-mail, uma vez que, “Do ponto de vista da standardização das perguntas e do potencial para transcrever as respostas os instrumentos [questionários] distribuídos por meio de e-mail têm grande potencial” (GÜNTHER, 1999, p.255).

Já para identificar as necessidades de otimização dos processos de planejamento e operacionalização da estratégia de busca para uma efetiva recuperação da informação pelos clientes/usuários (Alunos da Pós-Graduação) (**etapa c** desta pesquisa), utilizou-se a **observação sistemática** de seus processos de busca (escolhas das bases, utilização dos recursos, tempo gasto, etc) mediante a utilização de instrumento de observação (formulário) como guia (apêndice 2). A Observação, segundo Bains (1997) pode ter um papel útil em pesquisas de bibliotecas, especialmente se usado junto com outros métodos, como os questionários.

Uma amostra representativa dos sujeitos, foi convidada a elaborar um levantamento bibliográfico sobre o tema de suas próprias pesquisas. O histórico das buscas, contendo as estratégias formuladas pelos sujeitos, bem como o número de referências recuperadas será analisado com apoio bibliográfico.

Assim, para a realização da observação, de forma mais concisa possível, optou-se por extrair 3,5 % do total da população. O gráfico seguinte ilustra o número de sujeitos pertencente a cada estrato da população participante da pesquisa .





**Gráfico 2:** Número de participantes por amostragem estratificada proporcional para realização de observação

A observação foi realizada onde o usuário costuma fazer suas buscas quando esta nas dependências da UNESP-Marília, com dia e hora marcada, mediante a disponibilidade de tempo do usuário. Assim sendo, os locais utilizados foram o Laboratório de Informática destinado a todos os alunos do Campus e a própria Biblioteca.

É importante ressaltar que esta segunda amostra que foi retirada é uma amostra não probabilística (não se utilizou o sorteio) e apesar dos procedimentos não terem sido aleatórios, a amostra representa bem a população de onde foi extraída, pois está dividida em estratos. É ainda uma amostragem por conveniência, isto é, aqueles elementos que estiveram dispostos a colaborar e que se apresentaram, para a realização da observação de seus procedimentos de busca, fizeram parte da amostra que corresponde a 3,5% da população total, ou seja, 8 elementos.

Desta forma pretendeu-se, ao final desta pesquisa, oferecer subsídios à comunidade bibliotecária para que esta, utilizando-se das tecnologias de informação e comunicação atualmente empregadas, otimize e dinamize seus serviços de busca e referência, habilitando seus clientes/usuários para uma melhor efetividade da recuperação da informação, principalmente a especializada.

O próximo capítulo apresenta a análise dos resultados da pesquisa.

## CAPÍTULO 6

O silêncio é a atitude fundamental para se olhar e conhecer o mundo.

CYTRYNOWICZ<sup>41</sup>

## APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

### 6.1 Infra-estrutura disponível para pesquisa bibliográfica

Segundo dados obtidos junto à Seção Técnica de Referência, Atendimento ao Usuário e Documentação (**STRAUD**) da biblioteca da unidade (UNESP-FFC- Marília), atualmente os usuários têm acesso a todas as bases de dados assinadas pelo Conselho de Reitores das Universidades Estaduais paulistas (UNESP, USP e UNICAMP) – **CRUESP**.

De acordo com os dados obtidos no Site, o Consórcio CRUESP de Bibliotecas começou suas atividades em 1999, como um Grupo de Estudo, estabelecido pela resolução do Conselho de Reitores das Universidades do Estado de São Paulo (CRUESP<sup>42</sup>), apontando a integração dos Sistemas de Bibliotecas da USP (Universidade de São Paulo), UNESP (Universidade Estadual Paulista "Júlio de Mesquita Filho") e UNICAMP (Universidade Estadual de Campinas). O Consórcio teve por finalidade a instalação de bases de dados referenciais e de textos completos de revistas científicas internacionais, ambas eletrônicas,

---

<sup>41</sup> CYTRYNOWICZ, Roney. **Guerra sem guerra**: a mobilização e o cotidiano em São Paulo durante a segunda guerra mundial. São Paulo: Geração Editorial: Editora da Universidade de São Paulo (edusp), 2000.

<sup>42</sup> O Conselho de reitores das Universidades Estaduais de São Paulo - CRUESP - é constituído pelos reitores da USP, Unicamp e Unesp e pelos Secretários de Ciência e Tecnologia e da Educação. Tem como principais objetivos: fortalecer a interação entre as Universidades, propor possíveis formas de ação conjunta, conjugar esforços com vistas ao desenvolvimento das universidades, assessorar o Governador em assuntos de ensino superior, analisar e propor soluções para as questões relacionadas com o ensino e pesquisa nas Universidades Estaduais. O Decreto Estadual 26.914, de 15 de Março de 1987, dispõe sobre o Conselho de Reitores das Universidades Estaduais. Disponível em: <http://www.cruesp.sp.gov.br/>. Acesso em: 7 nov. 2004.

oriundas da cessão em comodato feita pela FAPESP do antigo Programa Biblioteca Eletrônica - ProBE, instituído em 1999 por este órgão de fomento, de interesse das Instituições que o integram, fazendo-as disponíveis a cada uma destas, em igualdade de condições de acesso. Hoje, o Consócio CRUESP de bibliotecas é um consórcio que inclui 89 bibliotecas que servem a quase 180.000 usuários (estudantes, faculdades e pesquisadores), além de outros usuários da comunidade externa. As propriedades dos três Sistemas de Biblioteca Universitários incluem uma coleção de mais de 4.470.000 itens.



**Figura 14:** Portal dos Sistemas de Bibliotecas das Universidades Estaduais Paulistas – CRUESP Bibliotecas

**Fonte:** <http://143.107.250.66/bibliotecas/CRUESP.htm>, acesso em 25 dez. 2004.

O Portal dos Sistemas de Bibliotecas das Universidades Estaduais Paulistas (do CRUESP), contém informações sobre as atividades desenvolvidas pelo consórcio e oferece os seguintes produtos:

- **Unibibli WEB** – permite o acesso simultâneo aos bancos de dados bibliográficos Dedalus/USP, Acervus/UNICAMP e Athena/UNESP e, diversos periódicos, por uma interface de busca unificada ;



**Bancos de Dados Disponíveis**

Seleção	Web Site	Descrição dos bancos de dados	Conteúdo	Mais...
<input checked="" type="checkbox"/>		USP - Universidade do Estado de São Paulo (Catálogo Global)	Acadêmica	<a href="#">info</a>
<input checked="" type="checkbox"/>		UNICAMP - Universidade Estadual de Campinas (Catálogo Global)	Acadêmica	<a href="#">info</a>
<input checked="" type="checkbox"/>		UNESP - Universidade Estadual Paulista Júlio de Mesquita Filho (Catálogo Global)	Acadêmica	<a href="#">info</a>

Escolher Todos

Encontre:

como

Palavras

Autor

Título

Assunto

Data

Buscar

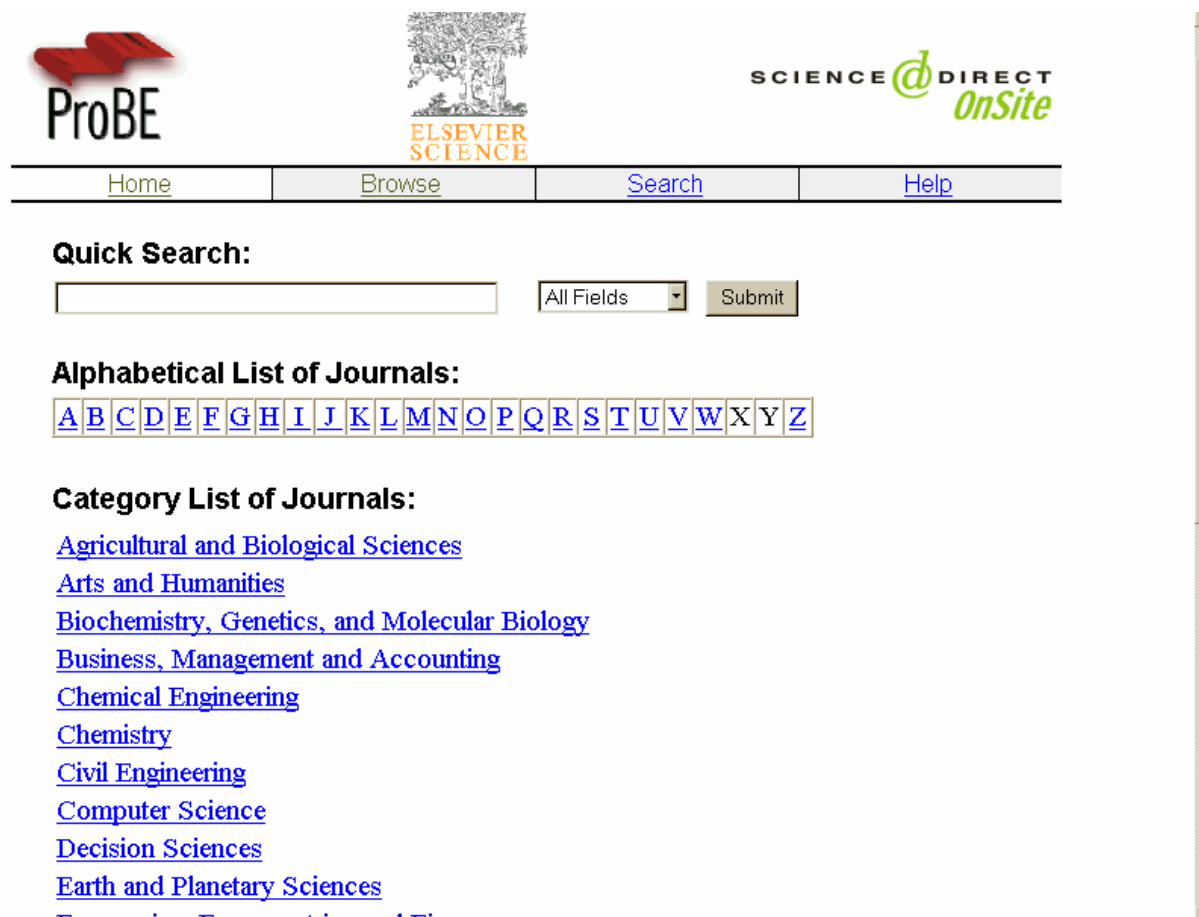
Copyright © 2002-2004 CRUESP Bibliotecas. (Acessos)

**Figura 15:** UnibibliWeb

**Fonte:** <http://www.cruesp.bc.unicamp.br/search.html>, acesso em 7 nov. 2004.

- **ERL** – Electronic Reference Library - (disponível em: <http://www.portaldapesquisa.com.br/databases/sites>) – 40 bases de dados de referência de assunto são disponibilizadas e acessadas, via Web, de qualquer computador ligado a uma das três redes de informação de Bibliotecas Universitárias, situado no posto de trabalho do pesquisador ou nas bibliotecas institucionais;

- **Biblioteca Eletrônica do CRUESP** - é formada pelo arquivamento de conteúdos digitais derivado de sociedades estabelecidas com editores comerciais e institucionais.



**Figura 16:** Biblioteca Eletrônica do CRUESP

**Fonte:** <http://probe.usp.br/cgi-bin/sciserv.pl?collection=journals>, acesso em 7 nov. 2004.

Outras bases são disponibilizadas no **Portal Bibliotecas UNESP**. Algumas de acesso restrito à intranet da UNESP (Web of Science, Current Contents, EbscoOnline, Swets, Computer Abstracts) e outras com acesso gratuito (SciELO – do Brasil e de vários países, IBICT, que podem ser acessadas via Internet).

**Portal Bibliotecas unesp**

Coordenadoria Geral de Bibliotecas / Rede de Bibliotecas / Catálogos Athena / Intranet / Fale com a CGB

**Banco de Dados Bibliográficos Athena**

Digitar palavra ou frase

Índice a percorrer:

Autor

Pesquisar Limpar

Clique aqui para outros tipos de busca no Athena !!!

**Portal Periódicos CAPES**

**Bibliotecas do CRUESP responsáveis por 59,44% dos acessos**

Recursos eletrônicos foram amplamente utilizados pela comunidade acadêmica do CRUESP

8.521.964 acessos no 1º semestre de 2004

**Ver Estatísticas**

**Biblioteca Digital da UNESP**

- Apresentação
- Teses On-Line
- Trabalhos publicados pela Rede de Bibliotecas
- Outras Bibliotecas Digitais

**Bases de Dados Digitais**

- CRUESP
- Periódicos On-Line
- Bases de Dados

**Serviços e Produtos**

- Convênios
- Contratos de Licença de Uso
- Cooperação
- Normalização Documentária para a Produção Científica da UNESP
- Scadunesp

**Destaque! Acesso à Safari Books Online**

**Figura 17:** Portal Bibliotecas UNESP

**Fonte:** <http://www.biblioteca.unesp.br>, acesso em 7 nov. 2004.

Os usuários têm também acesso ao **Portal de Periódicos** da Coordenação de Aperfeiçoamento de Pessoal de nível Superior – **CAPES**, que disponibiliza bases em todas as áreas do conhecimento.





**Figura 18:** Portal de Periódicos da CAPES

**Fonte:** <http://www.periodicos.capes.gov.br/> acesso em: 7 nov. 2004

O **Portal da Pesquisa** consiste em um dos produtos oferecidos pela **DotLib**, empresa especializada em prover acesso a bases de dados, que dentre seus clientes encontram-se grandes empresas, instituições de ensino superior e centros de pesquisa, tais como Petrobrás, Empresa Brasileira de Pesquisa Agropecuária - EMBRAPA, Coordenação de Aperfeiçoamento do Pessoal de Ensino Superior - CAPES, Universidade de São Paulo - USP, Universidade Estadual Paulista - UNESP, Universidade Estadual de Campinas - UNICAMP, Universidade Federal do Rio de Janeiro - UFRJ e Pontifícia Universidade Católica do Rio de Janeiro - PUC.



**Figura 19:** Portal da Pesquisa

**Fonte:** <http://www.portaldapesquisa.com.br/databases/sites>, acesso em 13 dez. 2004.

Dos portais citados acima, segundo dados obtidos na seção de referência, é possível adquirir o texto integral, através das bases de dados oferecidas, como a Inis, Scielo, Probe, EbscoOnline, SwetsWise e também através dos editores dos diversos periódicos oferecidos pela CAPES.

Quanto à acessibilidade, as bases do Portal da Pesquisa e da CAPES são de acesso restrito, sendo possível o acesso ao texto integral somente pelos computadores localizados dentro dos campus da Universidade, devido ao convênio desta com as instituições provedoras das bases de dados. Porém, foram disponibilizadas aos docentes e aos alunos da pós-graduação indicados pelos docentes, senhas para acesso fora da UNESP, mas somente ao Portal da Pesquisa (bases do CRUESP).



Segundo a opinião da Bibliotecária da seção de referência da referida unidade, as bases de dados mais utilizadas, pelos alunos da pós-graduação (são as bases específicas dentro de cada área), pois atendem as necessidades inerentes dos respectivos cursos: ERIC (Educação), Philosopher's Index (Filosofia), Sociological Abstracts (Ciências Sociais), LISA (Ciência da Informação) que são bases de dados referenciais, além das bases textuais de textos completos e os próprios periódicos eletrônicos.

Quanto ao o número de terminais que a biblioteca oferece para buscas on-line, a resposta foi a seguinte:

Infelizmente, a biblioteca está com uma defasagem de equipamentos. Os micros disponíveis são para a consulta dos materiais do acervo [local ou da rede UNESP de Bibliotecas]. Quando um usuário precisa fazer um levantamento, disponibilizamos um dos micros que possuímos na administração que são também utilizados pelos nossos estagiários. A biblioteca está insistindo firmemente junto à Direção visando a aquisição de novos equipamentos.<sup>43</sup>

Importante salientar que a UNESP-FFC-Marília, oferece também um laboratório de informática com 32 computadores para o acesso a Internet. Porém, este é utilizado por todos os alunos do campus (graduandos e pós-graduandos) de todos os cursos, mediante prévio agendamento e com um limite pré-estabelecido de horas por mês.

No que diz respeito a levantamentos bibliográficos em bases de dados, os tipos de serviços que são oferecidos aos usuários são:

---

<sup>43</sup> Informação obtida junto ao setor de referência da biblioteca da Unesp-FFC-Marília em setembro de 2004.

- Programa de Capacitação de Usuários em bases de dados (criado em 2002), e que a partir de 2003, foi dada uma maior ênfase, com cursos quinzenais (8 horas) para todos os usuários. Além desse horário, a critério do professor, a biblioteca oferece o curso em dias e horários a serem agendados para a classe toda.
- Também é disponibilizado aos usuários, mediante agendamento, o serviço de levantamento bibliográfico em bases de dados, com ou sem a presença do pesquisador (professor/aluno), para aqueles usuários que por algum motivo não podem comparecer aos cursos de capacitação. Atualmente, o setor de referência está fazendo o agendamento junto às Seções Administrativas e Departamentos do campus, visando a capacitação dos funcionários para o manuseio da base ATHENA.

Como foi aqui explanado, diversos recursos são atualmente disponibilizados aos usuários, para a busca em bases de dados nas mais diversas áreas do conhecimento. Há ainda as iniciativas para o livre acesso às teses e dissertações, além de outros materiais como a,

**Biblioteca Digital da UNICAMP**, [www.libdigi.unicamp](http://www.libdigi.unicamp) com um acervo digital significativo em teses e dissertações. **O Portal do Conhecimento da USP**, [www.saber.usp.br](http://www.saber.usp.br), disponibilizando a Biblioteca Digital de Obras Raras e Especiais da USP, e a Biblioteca Digital de Teses e Dissertações. A UNESP com **Portal da Biblioteca Digital** reunindo o vasto conteúdo das 31 Bibliotecas depositárias da produção científica da UNESP em um único Portal (CAMARGO et al., 2004, grifo nosso).

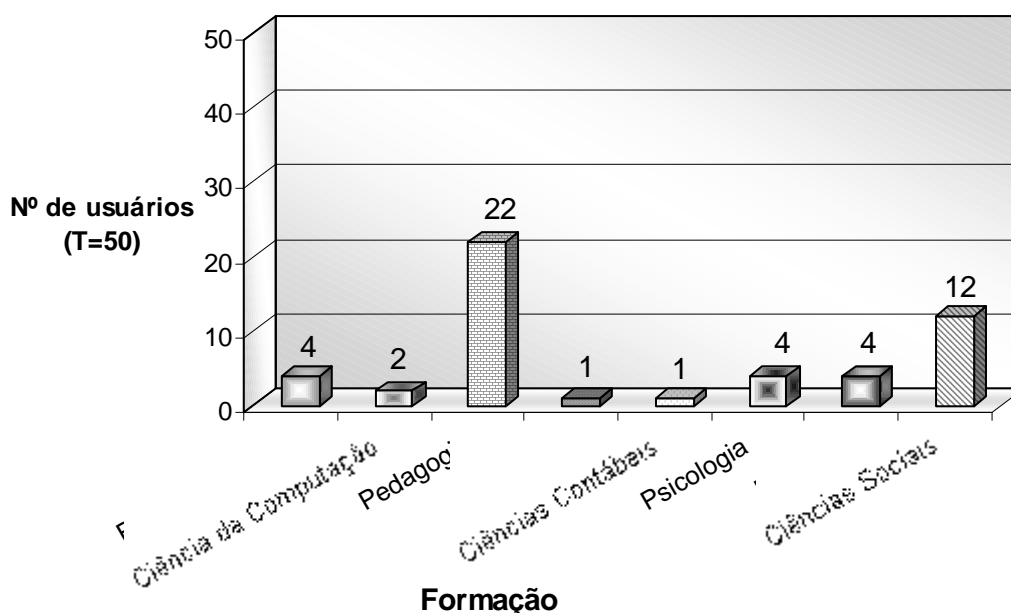
Isto possibilita o acesso, ao texto integral, de forma simultânea, a pesquisadores nacionais e estrangeiros, à produção científica destas instituições.

## 6.2 Necessidade informacional, conhecimentos e dificuldades na utilização das bases de dados pelos usuários

Buscou-se que se obtivesse 50 questionários respondidos. Assim, foi insistido que todos os participantes respondessem ao questionário. No caso da ausência de questionários respondidos – dos primeiros 50 questionários enviados, 13 (26%) não foram respondidos prontamente – Contactou-se outros alunos da pós-graduação para que respondessem ao questionário. Desta forma obteve-se 50 questionários respondidos que representam bem a população estudada.

O questionário objetivou levantar dados sobre: tipos de fontes utilizadas, formas de acesso a elas, grau de satisfação em relação às habilidades pessoais de busca e recuperação de informações, além dos conhecimentos e dificuldades na utilização das bases de dados.

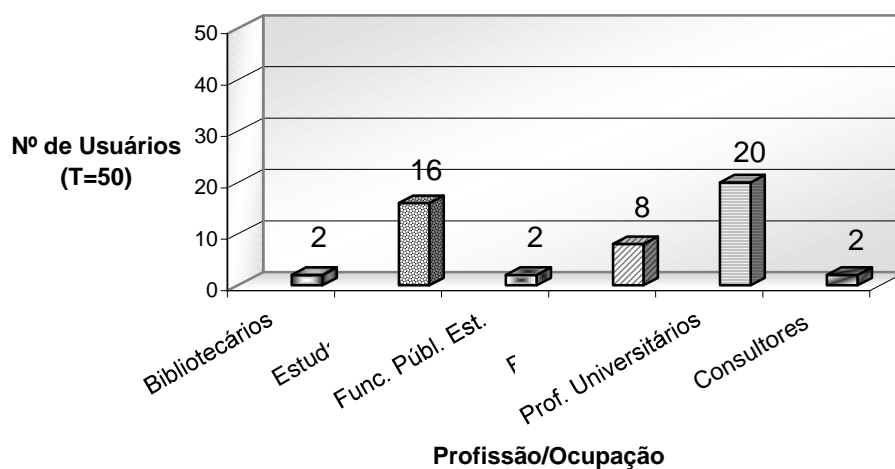
Os usuários foram questionados primeiramente quanto à formação acadêmica (gráfico 3) e atuação profissional (gráfico 4), dos quais os dados são apresentados abaixo;



**Gráfico 3:** Formação acadêmica dos alunos sujeitos da pesquisa

Sendo que os alunos advindos da Biblioteconomia e Ciência da Computação (6 alunos) pertencem ao programa de Pós-Graduação em Ciência da Informação; os advindos da Pedagogia, Direito, Ciências Contábeis e Psicologia (28 alunos) pertencem ao programa de Educação; os de Filosofia (4 alunos) e Ciências Sociais (12 alunos) pertencem aos respectivos programas de mesmo nome.

No que tange a profissão e/ou ocupação, pode-se ver no gráfico abaixo:



**Gráfico 4:** Profissão/ocupação dos alunos sujeitos da pesquisa

Percebe-se que 24 usuários (48%), aproximadamente metade da amostra, consideram-se apenas estudantes ou bolsistas, ou seja, pressupõe-se que tais alunos dediquem-se exclusivamente aos respectivos programas aos quais pertencem.

Os usuários foram questionados sobre qual local eles costumam realizar seus levantamentos bibliográficos. Pode-se ver as respostas na tabela que se segue:

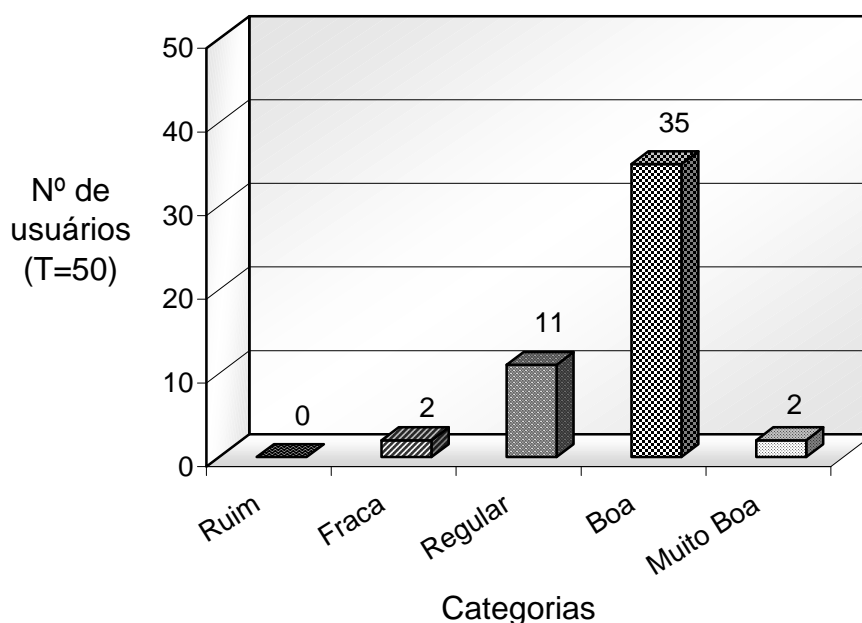
**Tabela 2 – Local onde o usuário costuma realizar buscas em bases de dados**

Local *	Frequência	%
Em casa	40	41,7
UNESP/Marília	37	38,5
No trabalho	13	13,5
Outro(s)	6	6,3
<b>Total</b>	<b>96</b>	<b>100</b>

\* múltipla escolha por 50 usuários.

De acordo com a **Tabela 2**, vários sujeitos indicaram realizar os seus levantamentos bibliográficos em mais de um local. A maior incidência (quarenta sujeitos ou 41,7%), no entanto, foi daqueles que realizam as suas pesquisas em suas residências, confirmando o que a literatura na área da biblioteconomia e Ciência da Informação vem apresentando em relação à autonomia e acesso que usuários finais cada vez mais estão adquirindo no que concerne a busca e recuperação da informação, mediante o avanço das tecnologias de informação e comunicação (BAINS, 1997; MERCADO, 1999; STURGES, 2001). Logo em seguida é citada a UNESP/Marília como local para a realização de buscas em bases de dados com trinta e sete (38,5%) usuários, o que indica que apesar da defasagem de equipamentos e da disponibilidade de muitas bases de dados estarem online e acessíveis de qualquer lugar, a biblioteca e/ou o laboratório de informática são utilizados pela maioria dos usuários para a busca nas bases de dados existentes para consulta. Treze sujeitos (13,5%) dizem utilizar o ambiente de trabalho para levantamentos bibliográficos e seis (6,3%) apontam outros locais como Cibercafés, bibliotecas de outros campus da UNESP e bibliotecas de outras instituições.

Foi solicitado aos participantes que opinassem em relação à infra-estrutura oferecida pela UNESP-FFC-Marília para acesso às bases de dados especializadas, os resultados estão indicados no gráfico a seguir:



**Gráfico 5:** Opinião dos usuários sobre a infra-estrutura oferecida

Apenas 4% (2 usuários) disseram que a infra-estrutura é **Muito Boa**, justificando que a biblioteca da autonomia ao pesquisador para a busca. A grande maioria, 70% (35 usuários), categorizou a infra-estrutura como sendo **Boa**, justificando há o apoio do pessoal da biblioteca e que o acervo local, sendo automatizado, agiliza a busca de informações. Porém, foram apontados pelos usuários que categorizaram como Boa a infra-estrutura, alguns pontos negativos, entre eles: poucos terminais para o acesso as bases de dados e melhorar a organização do acervo. 22% (11 usuários) categorizaram como **Regular** a infra-estrutura oferecida, apontando que; o laboratório e computadores são insuficientes para a demanda de alunos; problemas (não recebimento e/ou lentidão) de empréstimos entre bibliotecas, e o Comut; falta de funcionários; falta de informação sobre os recursos oferecidos; inexistência de uma base de dados de teses e dissertações; falta de indexação dos periódicos impressos disponíveis na biblioteca; atendimento via e-mail ineficiente; Tempo gasto elevado para a realização de buscas. Categorizaram como sendo **Fraca** 4% dos usuários consultados,

alegando que poucas vezes conseguiram localizar alguma coisa. Nenhum dos usuários (0%) disseram que a infra-estrutura oferecida é **Ruim**.

Dentro de uma realidade brasileira, como a de muitos outros países “em desenvolvimento”, “infelizmente o próprio acesso às tecnologias de informação e comunicação é muito deficiente [...]” (MCHOMBU; MENOU, 2004, p. 139) muitas vezes devido à “[...] limitações internas (por exemplo, falta de recursos, rigidez administrativa) bem como por limitações externas (por barreiras culturais ou pressões políticas) [...]” (MCHOMBU; MENOU, 2004, p. 147). Porém isto não justifica que os serviços prestados aos usuários sejam deficientes, pois nem a mais completa coleção, nem a melhor infra-estrutura em tecnologias da informação e da comunicação, poderá garantir a efetividade da recuperação da informação pelo usuário, se o bibliotecário falha ao administrar o processo de referência à luz do comportamento humano (FIGUEIREDO, 1999).

Segundo Figueiredo (1999, p. 110) “[...] a qualidade para um serviço de referência tem que ser auferida nos termos do atendimento das necessidades e das expectativas dos usuários, através de um comportamento adequado do corpo de pessoal e suas habilidades de comunicação [...]”. Isto porque,

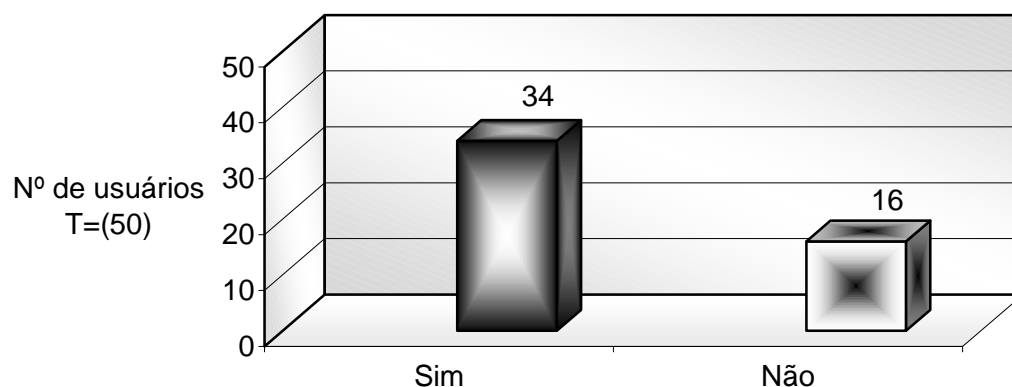
Cada biblioteca tem a sua clientela única, edifício e filosofia de serviço. Assim, cada programa de qualidade para o serviço de referência terá seus aspectos únicos. Nesta época de orçamentos apertados e justificativas para serviços, os bibliotecários devem atentar para algumas medidas tomadas na área de negócios/serviços. Se os administradores bibliotecários conhecem as necessidades e expectativas dos seus usuários, ficarão melhor equipados para oferecer qualidade, eficiência, serviço custo-benefício, com o pessoal e o orçamento disponíveis. Além disso, estarão construindo uma forte base



para apoiar argumentações para a manutenção ou o aumento de orçamento para a biblioteca (FIGUEIREDO, 1999, p. 116).

É preciso a criação de demandas, ou seja, fazer da referência um serviço de excelência que atenda satisfatoriamente os seus usuários para que, assim, se justifiquem os investimentos necessários.

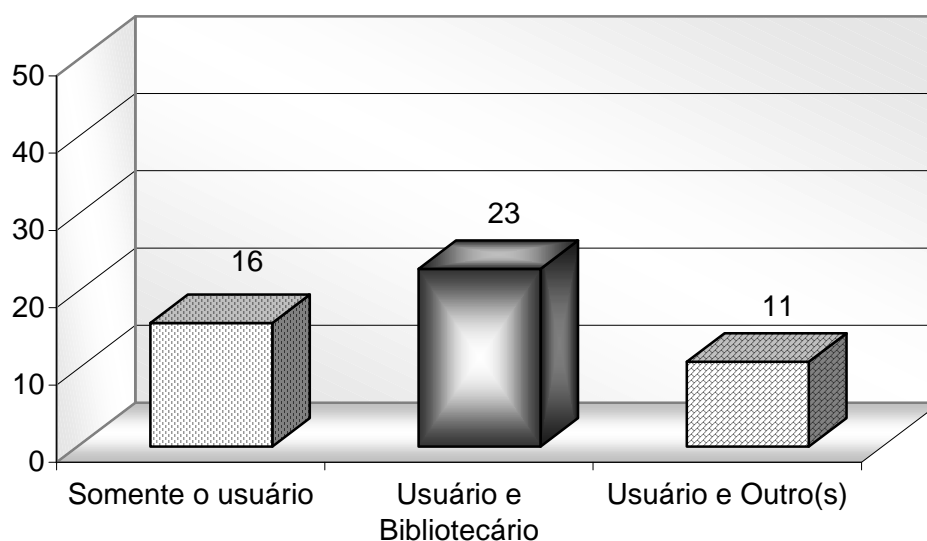
Em relação às dificuldades na realização de levantamentos bibliográficos em Bases de dados especializadas (**Gráfico 6**), 32% (16 usuários) alegaram não encontrar dificuldades, enquanto que 68% (34 usuários) admitiram ter algum tipo de dificuldade. Por mais que os sistemas de recuperação da informação tenham se tornados amigáveis, o usuário ainda enfrenta uma série de dificuldades para obter sucesso em suas buscas. Lopes (2002b, p.65) acrescenta que, “[...] em plena era do acesso à informação por meios eletrônicos, em que os usuários navegam na Internet com desenvoltura, [estes] ainda revelam as dificuldades inerentes ao processo de recuperação da informação em bases de dados.”



**Gráfico 6:** Dificuldades de busca em bases de dados especializadas

Ou seja, observa-se a partir dos dados coletados que, o processo de busca e recuperação dos usuários finais ainda é um fator de dificuldade que as tecnologias de informação e comunicação ainda não minimizaram.

Com relação a quem costuma fazer o levantamento bibliográfico para a realização de suas atividades acadêmicas (**Gráfico 7**), 32% (16 usuários) afirmaram realizar as próprias buscas sem o auxílio de terceiros, 46% (23 usuários) disseram que realizam buscas com o auxílio do bibliotecário e 22% (11 usuários) fazem buscas com a ajuda de outros, como: colegas, orientandos, estudantes de biblioteconomia, e outros.



**Gráfico 7:** Quem realiza levantamentos bibliográficos

Os motivos apontados para se recorrer a intermediários para a realização de levantamentos bibliográficos são bastante variados. Dos que afirmaram recorrer ao bibliotecário (46%) ou a outros intermediários (22%), na questão anterior (num total de 68%), para auxílio no levantamento bibliográfico, 18% afirmaram que o fazem para poupar tempo, enquanto 16% dizem ter desfamiliaridade em buscas em bases de dados. Outros motivos podem ser visto na tabela seguinte:

**Tabela 3- Motivos para a procura pelo bibliotecário ou outros para a realização de levantamentos bibliográficos**

<b>Motivos</b>	<b>Nº Usuários</b>	<b>%</b>
Para poupar tempo	9	18
Dificuldades na realização do processo de buscas	8	16
Dificuldades de recuperar documentos relevantes	6	12
Recorre ao Bibliotecário para garantir a precisão nas buscas	3	6
Dúvidas em como acessar as bases de dados	3	6
Falta de paciência para lidar com os sistemas de informação	2	4
Falta de capacidade para lidar com as mudanças tecnológicas	2	4
Acesso restrito que outros (colegas) tem	1	2
Não responderam (fazem buscas sozinhos)	16	32
<b>Total</b>	<b>50</b>	<b>100</b>

Pode-se notar que 12% dizem procurar o bibliotecário de referência ou outros por terem dificuldades em encontrar material sobre a temática de estudo, ou seja, têm dificuldade em recuperar itens relevantes. Outra resposta semelhante foi daqueles que afirmaram procurar o bibliotecário para garantir uma maior precisão nas buscas (três usuários ou 6%). Logo em seguida, dificuldades em como acessar as bases de dados (com 6% dos usuários) também é um dos motivos para a procura pelo bibliotecário ou outros intermediários, ou seja, eles sequer sabem como acessar as bases disponíveis. Interessante notar que a falta de paciência em lidar com sistemas de informação e a falta de capacidade para acompanhar as mudanças tecnológicas (ambas com 4%) também foram apontadas como um dos motivos. Sobre isto, Cuenca (1999, p. 294) afirma que “a dificuldade em lidar com as tecnologias sugere que o usuário deve ter alguma experiência prévia com o uso de computadores para conseguir maior sucesso nos resultados de suas buscas [...]”. Também foi apontado por um usuário (2%), que um dos motivos é o acesso restrito que outros (colegas) têm às bases de dados de outras instituições. Dezesesseis participantes (32%) não responderam esta questão, pois não recorrem a intermediários para a realização de seus levantamentos bibliográficos.

Apesar da alegação dos usuários de que a procura por intermediários para a recuperação da informação seja mais para poupar tempo, pode-se perceber que a unfamiliaridade com buscas em bases de dados é um item também relevante pois, como no estudo de Cuenca (1999, p. 293, grifo nosso), “[...] o fator principal foi a falta de familiaridade com as bases de dados justificada pelo fato de o bibliotecário [intermediário] conseguir **maior rapidez**, pertinência e eficácia nos resultados de busca”.

Assim, na medida em que, os usuários sozinhos encontram dificuldades nos processos de busca e recuperação da informação, Foster et al. (2002) afirma, pertinentemente, que a importância da interação entre usuário e intermediário bibliotecário como um meio para tornar a recuperação da informação possível esta clara. Durante a interação de busca pré-online, usuários e bibliotecários desenvolvem modelos de busca conjuntamente e compreendem parte dos problemas e objetivos do usuário, além de obter as respostas mais apropriadas do sistema (FOSTER et al, 2002).

Dentre as dificuldades apontadas em levantamentos bibliográficos em Bases de dados especializadas, pelos alunos da pós-graduação (**Tabela 4**), em questão de múltipla escolha, nota-se que a elaboração da estratégia de busca constou ser a principal dificuldade encontrada (38,2%). Pennanen e Vakkari (2003) já demonstraram em sua pesquisa que a especificidade da construção da estratégia de busca dos usuários finais, que estudaram, permaneceu totalmente inalterada, assim como, a habilidade para expressarem-se em linguagem de interrogação. Isto sugere a mesma idéia de Cardoso (2004, p. 5) quando diz que “[...] expressar uma necessidade de informação é uma tarefa difícil. Existe uma distância semântica entre a real necessidade dos usuários e o que ele expressa na consulta formulada”. Ou seja, traduzir a necessidade de informação para uma linguagem de interrogação que o sistema entenda, ainda é uma tarefa bastante complexa para os usuários finais.

**Tabela 4- Dificuldades encontradas pelos usuários em levantamentos bibliográficos em Bases de dados especializadas**

<b>Dificuldades*</b>	<b>Frequência</b>	<b>%</b>
Elaboração da estratégia de busca	29	38,2
Escolha da base de dados	20	26,3
Seleção dos termos (palavras-chave)	13	17,1
Falta de domínio do idioma das bases (inglês)	11	14,5
Outra(s)	3	3,9
<b>Total</b>	<b>76</b>	<b>100</b>

\* múltipla escolha por 50 usuários.

Em seguida vem a escolha da base de dados (26,3%) mostrando que também há dificuldades de localização das fontes de informação disponíveis. Isto indica a emergência de melhorias nas bibliotecas, de modo que facilite o processo de seleção de base de dados on-line, especialmente para o crescente número de usuários finais que têm acesso aos recursos da biblioteca remotamente (MA, 2002). Neste sentido, segundo Ma (2002, p. 569), muitas aplicações de sistemas especialistas (agentes e/ou ferramentas inteligentes) “[...] designados para ajudar usuários com a seleção de base de dados tem sido desenvolvidas [...]”<sup>44</sup>. Isto tem auxiliado o trabalho de referência no fornecimento das melhores fontes de informação para cada tipo de usuário. Ou seja, o processo de busca começa pela escolha daquela base de dados que poderá atender a uma necessidade de informação de forma satisfatória. Na mesma **Tabela 4** pode-se ver que a seleção de termos de busca, vem em terceiro lugar (17,1%) no que diz respeito a dificuldades encontradas, seguida de falta de domínio do idioma da base (14,5%) e outras dificuldades (3,9%) como o acesso restrito que algumas bases de dados têm ao serem acessadas fora dos domínios da UNESP.

Quanto às solicitações de algum tipo de auxílio ao bibliotecário de referência para a busca em bases de dados (**Tabela 5**), 36% dos usuários consultados apontaram a

<sup>44</sup> [...] designed to help users with database selection have been developed [...].

Escolha da base de dados como o principal motivo. Isto indica a dificuldade dos usuários de identificar a base de dados mais adequada que a instituição disponibiliza para consulta, e que poderá satisfazer uma necessidade de informação. Quanto a isso, Mercado (1999, p. 262) acrescenta que,

Bases de dados eletrônicas estão se tornando ainda mais sofisticadas e numerosas. Os estudantes, até mesmo quando suas habilidades com computador são boas, ainda necessitam de ajuda com a escolha da base de dados e com a avaliação de seus resultados de busca [...]<sup>45</sup>.

Isto é, “[...] nenhuma instituição, equipe de pesquisadores, estudante ou pós-graduado começará uma linha de pesquisa, tese, ou trabalho de pesquisa em geral, sem antes haver efetuado as correspondentes consultas nas bases de dados adequadas”<sup>46</sup> (MANUEL AGUADO, 1995, p. 22). Neste sentido, existe a emergência de guiar os usuários às fontes que melhor atenderão as suas necessidades informacionais.

A combinação de conectores de busca (AND; OR; NOT; etc), foi apontada por 26,7% dos usuários consultados, como motivo para a solicitação de auxílio ao bibliotecário de referência. “O desenvolvimento das estratégias de busca não é uma habilidade natural”<sup>47</sup> (TENOPIR, 2003, p. 632). Segundo Pennanen e Vakkari (2003) os usuários finais não são familiarizados com a Lógica Booleana. E Tenopir (2003) aponta ainda que os usuários finais,

---

<sup>45</sup> Electronic databases are becoming more sophisticated and numerous. Students, even when their computer skills are good, still need help with database choice and evaluation of their search results [...].

<sup>46</sup> [...] ninguna Institución, equipo de investigadores, estudiante o postgraduado comenzará una línea de investigación, tesis o trabajo de investigación en general, sin antes haber efectuado las correspondientes consultas a las bases de datos adecuadas.

<sup>47</sup> Developing search strategies is not a natural skill.

mesmo depois de receberem capacitação em busca em bases de dados, estes ainda necessitaram de bibliotecários para ajuda-los a construir e refinar buscas complicadas.

**Tabela 5 – Motivos para solicitações de auxílio ao bibliotecário de referência**

<b>Motivos*</b>	<b>Frequência</b>	<b>%</b>
Escolha das bases de dados	31	36
Combinar comandos (and; or; not; adj; etc)	23	26,7
Escolha de termos de busca (palavras-chave)	19	22,1
Não peço auxílio nenhum	11	12,8
Outro(s)	2	2,3
<b>Total</b>	<b>86</b>	<b>100</b>

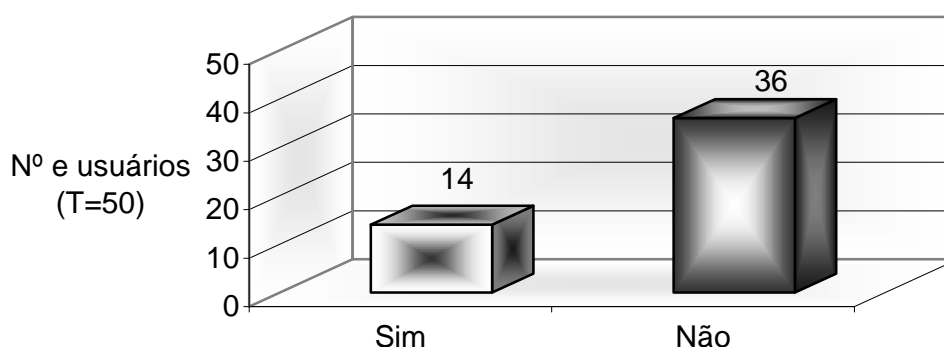
**\* múltipla escolha por 50 usuários.**

A escolha dos termos de busca foi motivo para 22,1% dos usuários para a procura pelo bibliotecário. Spink e Saracevic (1997) propõem que a seleção de termos de busca é um dos objetivos e processos fundamentais na interação da recuperação da informação. Eles apontam cinco fontes de termos de busca em um estudo que investigou a efetividade dos termos durante a busca online. Estas fontes são: o estabelecimento da questão; a interação do usuário com o intermediário; o Tesouro; o próprio intermediário; e a relevância de feedback dos termos. Foster et al (2002) em seu estudo sobre pesquisa mediada e busca da informação constatou que o processo de interação ajudou o usuário a obter resultados muito úteis devido à experiência do intermediário em definir os parâmetros de busca, refinando as respostas do sistema. Ou seja, a seleção de termos de busca para uma pergunta e, a construção de questões para serem apresentadas ao sistema de informação, é um processo altamente interativo, sendo uma importante parte do processo de interação. 12,8% dos usuários alegaram não pedir nenhum tipo de auxílio ao bibliotecário, enquanto que 2,3% dizem pedir auxílio quando desconhecem a abrangência (ou área) de uma determinada base de dados. Autores como Spink, Goodrum e Robins (1998) constataram que os intermediários são solicitados para

diferentes tipos de informações por seus usuários incluindo, solicitações de informações sobre seleção de base de dados, elaboração de estratégias de busca, seleção de termos de busca, etc.

Comparando a **Tabela 4** com a **Tabela 5**, percebe-se que, mesmo perante a principal dificuldade encontrada pelos usuários (elaboração da estratégia de busca), estes preferem solicitar auxílio ao bibliotecário principalmente para a escolha das bases de dados, e não para a elaboração de estratégias de buscas.

Quando solicitados a responder se já participaram de algum curso ou treinamento para o uso de bases de dados especializadas, os usuários responderam da seguinte maneira (**Gráfico 8**);



**Gráfico 8:** Participação em curso ou treinamento em bases de dados especializadas

Percebe-se que a participação em capacitação ou cursos para a utilização de bases de dados, pelos usuários estudados, é relativamente baixa (14 usuários ou 28%) enquanto que 72% (36 usuários) afirmaram nunca terem participado de nenhum tipo de curso para a capacitação em buscas em bases de dados.



**Tabela 6 – Local e Contribuições do curso ou treinamento para o uso de bases de dados especializadas**

<b>Contribuições</b>	<b>Local</b>	<b>Nº usuários</b>
Contribuições no entendimento sobre a lógica Booleana e na utilização de base de dados.	Unesp-FFC-Marília	3
Na Uel. esclareceu dúvidas, mas insuficiente para dominar as buscas	UEL	3
Aprendizagem de alguma coisa	Unesp-FFC-Marília	2
Contribuiu para pesquisa em bancos de dados, mas ainda tem dificuldades	Não respondeu	2
Conhecimento de bases de dados, usa-las corretamente	Não respondeu	2
Contribuições foram necessárias devido as mudanças ocorridas nas bases de dados (mudanças de interface, novos recursos etc)	Unesp CGB-SP	1
Melhora significativa no desempenho de levantamentos bibliográficos posteriormente ao treinamento	No trabalho	1
<b>Total</b>		<b>14</b>

Quanto ao local onde os respondentes realizaram esses cursos e/ou treinamentos, verificou-se que cinco participaram de cursos oferecidos pela Biblioteca da Unesp – Marília, três em cursos e /ou treinamentos oferecidos pela Universidade Estadual de Londrina - UEL, um afirmou ter participado de um curso oferecido pela Coordenadoria Geral de Bibliotecas da Unesp e um afirmou ter recebido este tipo de treinamento em seu próprio local de trabalho. Quatro respondentes não identificaram o local onde participaram deste tipo de atividade. Os outros 36 sujeitos não responderam esta questão, pois alegaram nunca ter participado de cursos de treinamento para uso de bases de dados.

Estes resultados demonstram que embora a biblioteca do campus ofereça cursos para utilização de bases de dados especializadas ainda há uma baixa participação dos alunos dos cursos de Pós-Graduação da casa, pois somente cinco dentre os 50 respondentes, afirmaram ter participado deste tipo de curso. Resta investigar através de outros estudos qual o motivo para esta baixa participação. Com relação aos demais locais onde os sujeitos realizaram treinamentos, verificamos que a UEL também tem procurado instruir os usuários para ao uso das bases. No entanto, faltam dados qualitativos a respeito dos cursos e ou

treinamentos realizados pelos sujeitos, como por exemplo, número de horas e recursos utilizados.

Quando questionados sobre as contribuições que o curso/treinamento ofereceu, a maioria respondeu que de, alguma forma, houve melhorias no entendimento da lógica booleana; no conhecimento sobre fontes de informação e como usa-las corretamente; porém não o suficiente para dominar a busca em bases de dados.

Considerando a baixa procura pela capacitação, pelos usuários finais de bases de dados especializadas e, também, por aqueles que participam dos cursos/treinamentos, o aproveitamento que estes obtém é parcial (pois ainda sim existem dificuldades), leva-se a crer que os usuários

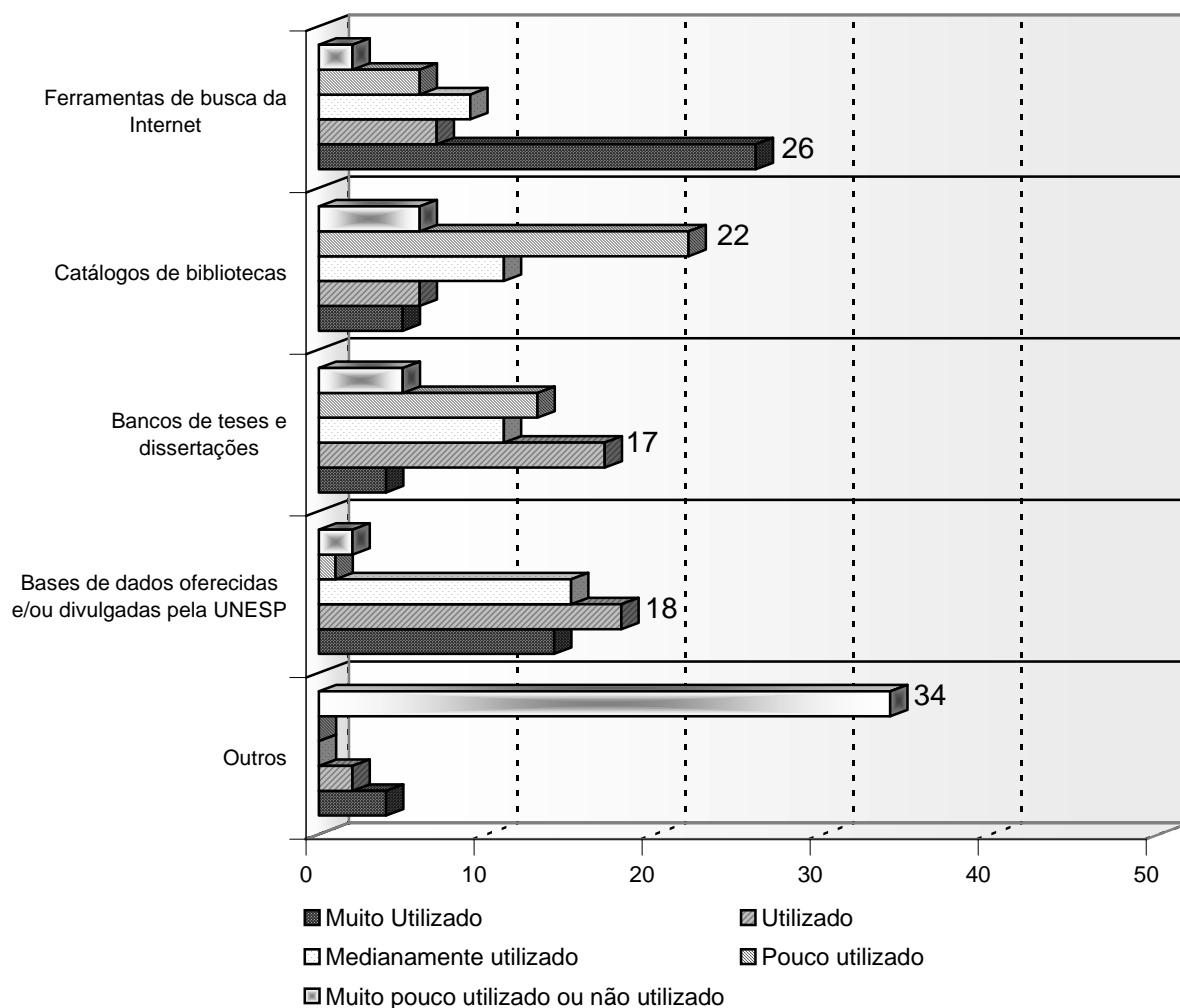
Não estão motivados para utilizar estas tecnologias, nem para aprenderem sobre seu uso. Esta motivação se daria a partir de perceber sua utilidade em suas atividades principais e as vantagens de seu uso, necessitam saber as potencialidades dos recursos para uma específica aplicação imediata [...] <sup>48</sup>  
(HERNÁNDEZ SALAZAR, 2003, p. 174).

É neste sentido que o papel do bibliotecário de referência deve ser focado, de forma ativa, para a importância e as potencialidades que as bases de dados atualmente oferecem aos usuários finais, seja através da capacitação, divulgação, ou serviço de referência. Entretanto, como será visto mais adiante há, cada vez mais, a emergência do bibliotecário em atuar com as tecnologias de informação e comunicação na interação com o usuário.

---

<sup>48</sup> No están motivados para utilizar estas tecnologías, ni para formarse sobre su uso. Esta motivación se daría a partir de percibir su utilidad en sus actividades principales y las ventajas de su uso, necesitan saber las potencialidades de los recursos para precisar una aplicación inmediata [...].

Foi solicitado aos usuários que assinalassem os tipos de fontes que costumam utilizar. Os dados obtidos geraram o gráfico que se segue:



**Gráfico 9:** Representação gráfica das fontes utilizadas na busca e recuperação de informações

O gráfico acima mostra as categorizações que os usuários deram a diferentes fontes de informação eletrônicas. Assim, 26 usuários (52%) categorizaram como **muito utilizadas**, as ferramentas de busca da Internet. Segundo Salvador Oliván e Angós Ullate (1999, p. 48),

Na Internet se pode utilizar fontes de informação públicas (geralmente gratuitas) ou privadas (pagando por sua utilização). Estas fontes incluem revistas eletrônicas, periódicos, diretórios, livros, arquivos de som, imagens, bases de dados científicas, de produtos e serviços, textos, jogos, etc.<sup>49</sup>

Ou seja, através de sistemas de busca da Internet, cada vez mais eficientes, pode-se acessar de forma relativamente fácil aos diversos recursos disponíveis na rede. O que pode-se justificar a grande intensidade da utilização das ferramentas de busca da Web em relação as demais fontes.

Já os catálogos de bibliotecas foram categorizados como **pouco utilizados** (por 44% dos usuários). Este resultado indica a baixa frequência com que os usuários visitam a biblioteca para buscas no catálogo local, assim como o acesso aos catálogos de bibliotecas de outras instituições. Bancos de teses e dissertações foram categorizadas como **utilizadas** por 34% dos usuários, assim como, as bases de dados oferecidas e/ou divulgadas pela UNESP foram por 36% dos usuários consultados. Um número de utilização muito próximo entre estas duas fontes apontando, pode-se dizer, para o crescimento da importância dos bancos de teses e dissertações na medida em que estas vêm se consolidando cada vez mais como uma das formas de divulgação da pesquisa científica acadêmica. A grande maioria (68%) **não utiliza** outras fontes para obter informações, apenas 9 alunos (18%) afirmaram utilizar **muito pouco** o orientador para a obtenção de informação, diferentemente do que constatou Caregnato (2003) em seu estudo quando mediu a relevância, a confiabilidade e o volume de informações que as fontes de informação avaliadas tem em relação aos usuários estudados, com professores e orientadores aparecendo em destaque.

---

<sup>49</sup> En Internet se pueden utilizar fuentes de información públicas (generalmente gratuitas) o privadas (pagando por su utilización). Estas fuentes incluyen revistas electrónicas, periódicos, directorios, libros, archivos de sonido, imágenes, bases de datos científicas, de productos y servicios, textos, juegos, etc.

Solicitou-se também, que os usuários assinalassem as bases de dados que são mais utilizadas seus para levantamentos bibliográficos. Com os dados coletados, obteve-se a seguinte tabela:

**Tabela 7 – Bases de dados mais utilizadas, segundo nº de usuários**

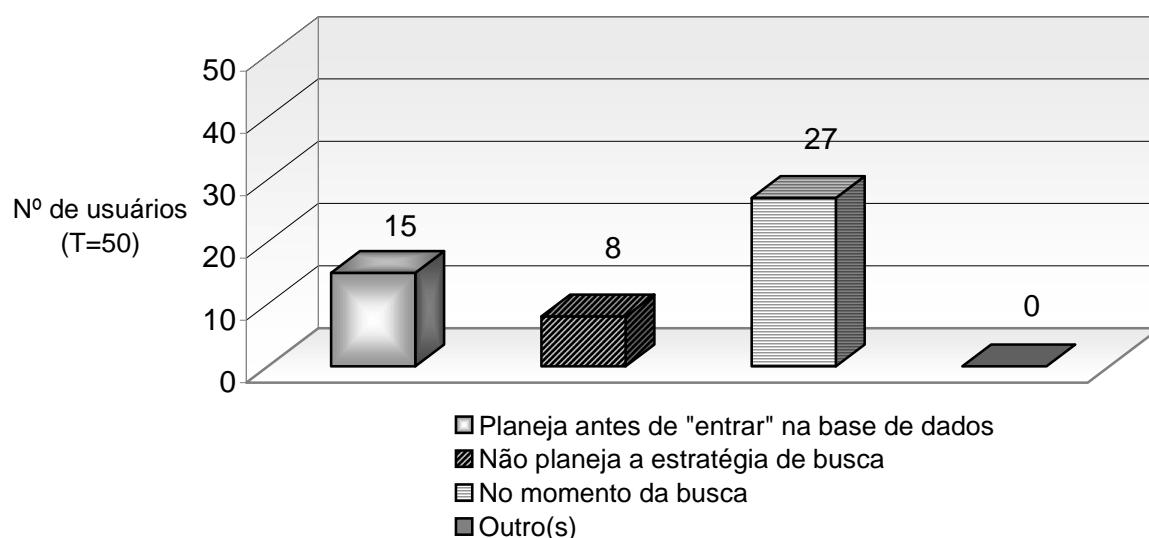
Bases de dados *	Frequência	%
CAPES	38	16,4
SciELO	33	14,3
ATHENA	29	12,5
Dédalus	21	9,1
Web of Science	17	7,4
Eric	16	7
Unibibli	13	5,6
IBICT	10	4,3
Cruesp bibliotecas	9	3,9
Ebsco Online	7	3
British Library	6	2,6
Library of Congress	6	2,6
Prossiga	6	2,6
Probe	5	2,2
Current Contents	5	2,2
Acervus	4	1,7
Emerald-Library	3	1,3
Outra(s)	3	1,3
Swets	0	0
DBFCC	0	0
Total	231	100

\* múltipla escolha por 50 usuários.

Conforme se pode ver na tabela acima, 16,4% dos usuários consultados elegeram o portal de pesquisa **CAPES** como o recurso mais utilizado para a busca em bases de dados. Logo após, vem a **SciELO** com 14,3%, seguida da base de dados catalográfica **Athena** (pertencente a UNESP) com 12,5% das indicações e a base de dados catalográfica **Dédalus**

com 9,1%. A utilização de outras bases pode ser observada conforme se apresenta na **Tabela 7**. De acordo com os dados obtidos, os usuários utilizam as bases de dados disponíveis para consulta (algumas mais outras menos) para realizarem seus levantamentos bibliográficos.

Quanto ao planejamento da Estratégia de busca (**Gráfico 10**), 15 usuários (30%) disseram que o fazem antes de entrar na base de dados, enquanto que 27 usuários (54%) planeja a estratégia no momento da busca. Dos usuários consultados, 8 (16%) afirmaram não planejar a estratégia de busca.



**Gráfico 10:** Planejamento da estratégia de busca

Conforme o estudo realizado por Cuenca (1999, p. 294), no preparo de estratégia de busca pelo usuário com autonomia no uso das bases, foi verificado também “que poucos preparam sua estratégia para realizar busca”.

Segundo Bertholino (1999) a formulação da estratégia de busca é fundamental para refinar a busca e poder obter resultados relevantes aos interesses do usuário. É importante, na definição das palavras-chave, informar os termos sinônimos, correlacionados e equivalentes, bem como suas respectivas definições no idioma inglês, adotado pela maioria

das bases internacionais, e também consultar tesouros, vocabulários controlados, dicionários especializados e outras fontes.

Há ainda a necessidade de saber como e quando expandir a questão (em linguagem de interrogação) apresentada ao sistema de recuperação da informação para aumentar, através dos termos de busca adicionais e da lógica booleana aplicada apropriadamente, o desempenho de recuperação, assim como, utilizar-se do feedback do sistema, pois, a expansão da questão, baseada na avaliação de relevância (feedback) é um modo importante de prover os pesquisadores com termos de busca extraídos de itens com potenciais informações pertinentes (PENNANEN; VAKKARI, 2003).

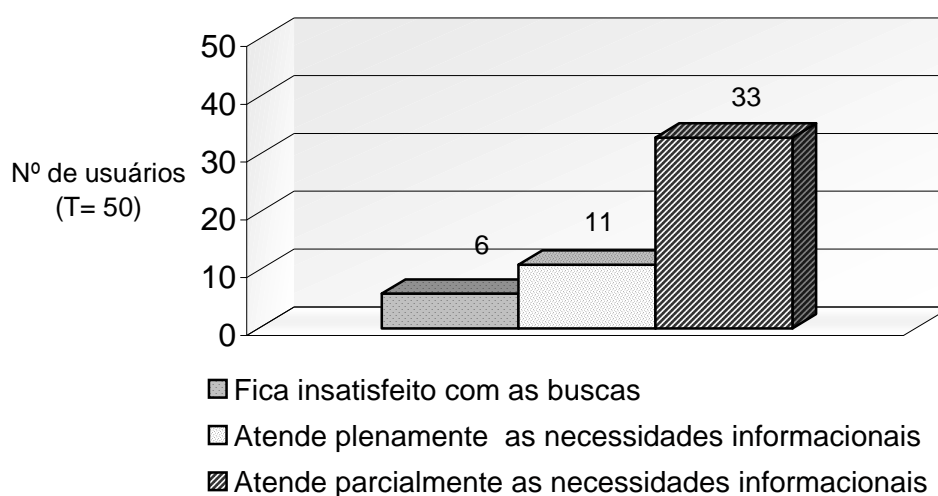
Quanto a isso (**Tabela 8**), 66,7% dos usuários, afirmaram que reformulam a sua estratégia de busca para a obtenção de melhores resultados, enquanto que 19% preferem mudar de base de dados com a mesma estratégia de busca. 8% preferem entrar em contato com o orientador ou mudar de base de dados com outras estratégias de busca e 6,3% se conformam com os resultados obtidos.

**Tabela 8 – Procedimentos adotados pelos usuários quando os resultados são insatisfatórios na busca em base de dados**

<b>Procedimento adotado *</b>	<b>Frequência</b>	<b>%</b>
Reformula sua estratégia de busca	42	66,7
Muda de base de dados com a mesma estratégia de busca	12	19
Outro(s)	5	8
Conforma-se com o resultado obtido	4	6,3
<b>Total</b>	<b>63</b>	<b>100</b>

**\* múltipla escolha por 50 usuários.**

Foi solicitado aos alunos da pós-graduação em estudo, que apontassem o grau de satisfação em relação às próprias buscas, conforme é visto no gráfico que se segue:



**Gráfico 11:** Satisfação dos usuários quanto às próprias buscas

A maioria dos respondentes, 66% (33 usuários) assinalou que as próprias buscas atendem parcial as suas necessidades informacionais, enquanto que 22% (11 usuários) disseram que se sentem satisfeitos com as suas próprias buscas. Ficam insatisfeitos com as próprias buscas, 12% dos usuários participantes da pesquisa.

Este resultado, contradiz o que foi constatado na pesquisa de Caregnato (2003), que estudou o comportamento de busca de informações de alunos do programa de Pós-Graduação em Comunicação e Informação da UFRGS. Segundo seu estudo, a maioria dos alunos, sentem-se satisfeitos ou muito satisfeitos (71%) com as habilidades pessoais de busca e uso de informações. Isto pode apontar para as necessidades específicas que cada comunidade de usuários e, por que não dizer, as necessidades específicas e individuais de informação que cada usuário possui.

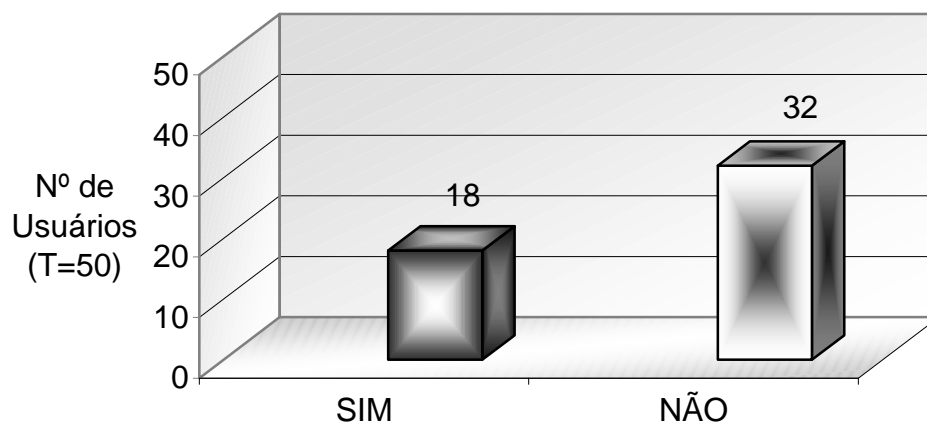
Para Cuenca (1999, p. 294), “é sabido que uma boa estratégia de busca assegura resultados mais relevantes de citações recuperadas. A insatisfação com os resultados das buscas pode estar relacionada ao pouco tempo que o usuário final dedica ao preparo de sua



estratégia de busca”. O que foi visto no Gráfico 10, onde a maioria decide sua estratégia no momento da busca.

Porém, o conhecimento sobre os mecanismos de auxílio a busca é também importante para um bom resultado de busca. Saber o funcionamento de cada base de dados, os recursos disponíveis, como Tesouros, possibilidade de truncamento de raízes de palavras, operadores de proximidade de termos de busca (além da lógica booleana), possibilidade de busca por diversos campos (autor, título, assunto), ou no texto completo, delimitação de tempo, entre outros, possibilita otimizar as buscas em bases de dados.

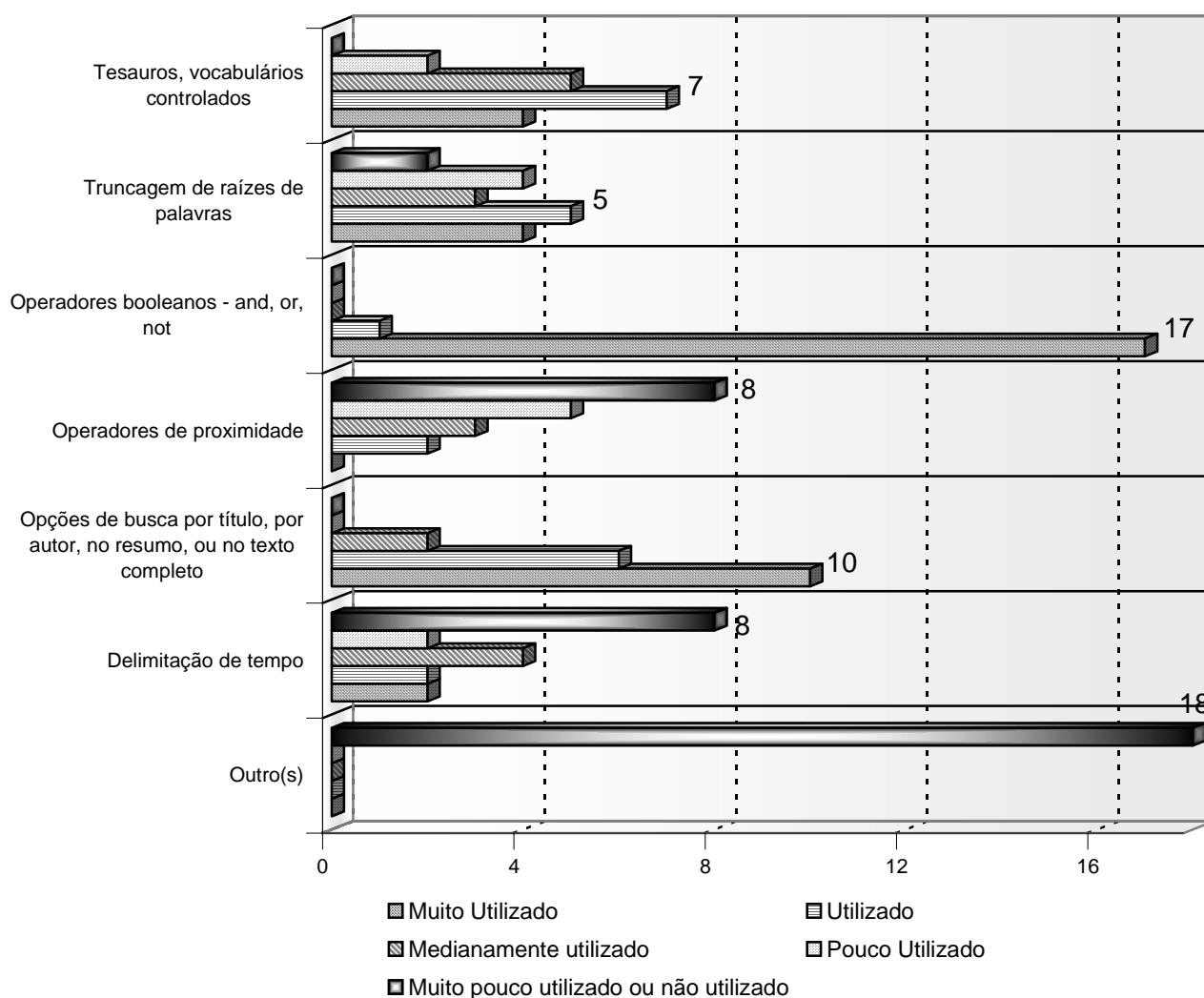
No que concerne sobre o conhecimento desses mecanismos, conforme o **Gráfico 12**, apenas 36% (18 usuários) afirmaram ter alguma familiaridade com os instrumentos de auxílio à busca, enquanto que 64% (32 usuários) dizem não possuir o conhecimento para utilizar tais recursos.



**Gráfico 12:** Familiaridade com os instrumentos de auxílio à busca

Ou seja, muitos usuários não utilizam estratégias de busca mais complexas. Conforme Tasso et al. (2002) vários estudos mostraram que os usuários frequentemente encontram dificuldades por causa de suas pobres habilidades para implementar estratégias de busca efetivas.

Segundo Tasso et al (2002), os usuários constroem questões, em linguagem de interrogação, usando somente termos que eles têm na mente, sem consultar os tesauros disponíveis, não levando vantagem da interação com os bancos de dados; eles executam somente pequenas modificações da questão e refinamentos. Além disso, utilizam-se principalmente da aplicação de *screen-browsing* (“folheio de telas”) e estratégias de busca na tentativa e erro (as duas menos efetivas estratégias), considerando que a maioria dos bibliotecários de referência e pesquisadores sofisticados adotam as heurísticas opções de busca e o uso estratégico dos tesauros (TASSO, et al.2002).



**Gráfico 13:** Utilização dos Instrumentos de auxílio a busca em bases de dados

Conforme pode-se perceber no **Gráfico 13**, dos 18 usuários (36% do total) que afirmaram ter alguma familiaridade com os instrumentos de auxílio a busca, 94% (17 usuários) categorizaram como **muito utilizado** os operadores booleanos (AND; OR; NOT), na busca em bases de dados. 55,5% (10 usuários) também definiram como **muito utilizado** as ferramentas de opções de busca (por título; autor; no resumo; texto completo). Operadores de proximidade, e a opção por delimitação de tempo (para buscas de documentos de um determinado período), foram categorizado por 8 usuários (44%) como **muito pouco utilizado ou não utilizado**, enquanto que, os tesauros e a truncagem das raízes das palavras foram consideradas como **utilizadas**, respectivamente, por 7 (39%) e 5 (27,8%) dos usuários respondentes desta questão. Nenhum dos usuários citou outro tipo de instrumentos de auxílio à busca.

Apesar do baixo índice de respondentes, esta questão mostrou que usuários um pouco mais familiarizados com buscas em bases de dados especializadas, são capazes de utilizar alguns dos instrumentos complementares de auxílio à busca. Mas, conforme destaca Tasso (2002, p. 344), “[...] até mesmo usuários experientes podem encontrar dificuldades para aplicar estratégias efetivas e, embora construam questões [de busca] mais complexas e façam uso do tesouro, o desempenho deles pode ser tão pobre quanto o de usuários inexperientes”<sup>50</sup>.

Aqui levantou-se uma série de dados destacando os diversos problemas que envolvem a recuperação da informação em bases de dados e a necessidade das interações entre o usuário, o bibliotecário e os sistemas de recuperação da informação atualmente disponíveis. A seguir, os dados obtidos durante a observação, complementarà a análise.

---

<sup>50</sup> [...] even experienced users may find it difficult to apply effective strategies and, although they construct more complex queries and make use of thesauri, their performance may be as poor as that of novices.

### **6.3 Necessidade de otimização dos processos de planejamento e operacionalização da estratégia de busca em bases de dados**

Nesta etapa da pesquisa, objetivou-se observar os processos de planejamento e operacionalização das estratégias de busca (seleção de termos de busca, seleção de base de dados, tipos de campos de busca utilizados, número de registros visualizados, número de registros relevantes recuperados, tempo gasto na busca), de forma a verificar como os sujeitos realizam buscas, para o diagnóstico de problemas e indicar as necessidades de otimização dos processos de busca em bases de dados especializadas.

Os sujeitos aqui observados (no total de 8 participantes) foram convidados a participarem, não sendo necessariamente os mesmos que participaram respondendo o questionário.

Assim sendo, mediante instrumento para a coleta dos dados, segue as observações realizadas:

**Caracterização e tema da busca do Usuário observado 1:**

<b>Pós-Graduação em</b>	Ciência da Informação
<b>Formação acadêmica</b>	Análise de Sistemas
<b>Profissão</b>	Analista de Sistemas
<b>Tema da busca</b>	Audiovisual
<b>Objetivo da Busca</b>	Trabalho p/ a Disciplina Elementos de Organização da Informação de Documentos Audiovisuais

O usuário observado 1 realizou duas sessões de busca as quais estão demonstradas a seguir:

(Busca 1- Usuário1)

<b>Procedimento para escolha da base de dados</b>	Endereço eletrônico		
<b>Base de dados</b>	GOOGLE – <a href="http://www.google.com.br">http://www.google.com.br</a> (ferramenta de busca da WEB) .		
<b>Reconhece facilmente o campo de busca</b>	SIM		
<b>Fonte dos descritores (termos de busca)</b>	Aula profº Jorge Caldeira Serrano		
<b>Pontos de acesso utilizados</b>	Todos que a ferramenta de busca permite (Tags)		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
1) BANCO DE DADOS AUDIOVISUAIS	simples	8830	5
2) ARQUIVOS AUDIOVISUAIS	avançada (docs. em português e em PDF)	892	9
3) ARQUIVAMENTO DIGITAL	simples	13700	4
4) ARQUIVO AUDIOVISUAL E TELEVISÃO	Avançada (uso do conector “E”)	4860	8
<b>Reações e/ou emoções do usuário durante a busca</b>	“geralmente links em PDF são Trabalhos acadêmicos”		
<b>Anotações do observador</b>	Utiliza mais o Google; Vê o resumo gerado e no máximo até 4º página Web de resultados que a ferramenta de busca gera; Estratégias feitas na hora		

(Busca 2 – Usuário 1)

<b>Procedimento para escolha da base de dados</b>	Ícone na área de Trabalho		
<b>Base de dados</b>	Athena – <a href="http://200.145.174.8:4505/ALEPH">http://200.145.174.8:4505/ALEPH</a>		
<b>Reconhece facilmente o campo de busca</b>	SIM		
<b>Fonte dos descritores (termos de busca)</b>	Aula profº Jorge Caldeira Serrano		
<b>Pontos de acesso utilizados</b>	1º Todos que a busca simples utiliza; 2º Assunto		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
5) ARQUIVOS AUDIOVISUAIS	simples	111	10
6) ARMAZENAMENTO AUDIOVISUAL	avançada (Assunto)	0	--
7) ARMAZENAMENTO AUDIOVISUAL	simples	32	6
<b>Reações e/ou emoções do usuário durante a busca</b>			
<b>Anotações do observador</b>	Estratégias feitas na hora		

**Tempo total de busca (usuário 1):** 00: 31 min.

**Caracterização e tema da busca do Usuário observado 2:**

<b>Pós-Graduação em</b>	Ciências sociais
<b>Formação acadêmica</b>	Ciências sociais
<b>Profissão</b>	Estudante
<b>Tema da busca</b>	Formas de organização da classe operária (apreciação histórica)
<b>Objetivo da Busca</b>	Trabalho p/ a Disciplina do curso

(Busca 1- Usuário 2)

<b>Procedimento para escolha da base de dados</b>	Ícone na área de Trabalho		
<b>Base de dados</b>	Athena – <a href="http://200.145.174.8:4505/ALEPH">http://200.145.174.8:4505/ALEPH</a>		
<b>Reconhece facilmente o campo de busca</b>	SIM		
<b>Fonte dos descritores (termos de busca)</b>	Memória		
<b>Pontos de acesso utilizados</b>	1º Todos que a busca simples utiliza; 2º Título		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
1) ORGANIZAÇÃO CLASSE OPERÁRIA	simples	573	4
2) ORGANIZAÇÃO CLASSE OPERÁRIA HISTORIA	simples	5985	--
3) ORGANIZAÇÃO CLASSE OPERÁRIA HISTORIA LUTA POLÍTICA	simples	2912	--
4) CLASSE OPERÁRIA	avançada	7	1
5) CLASSE OPERÁRIA FORMAS	simples	0	0
<b>Reações e/ou emoções do usuário durante a busca</b>	“veio muita coisa”		
<b>Anotações do observador</b>	Estratégias feitas na hora; diz preferir os livros do acervo local; índice muito baixo de referências relevantes recuperadas; rever palavras-chave		

**Tempo total de busca (usuário 2):** 00: 27 min.

**Caracterização e tema da busca do Usuário observado 3:**

<b>Pós-Graduação em</b>	Ciências sociais
<b>Formação acadêmica</b>	Ciências sociais
<b>Profissão</b>	Estudante
<b>Tema da busca</b>	Tecnologia e cultura
<b>Objetivo da Busca</b>	Trabalho para a disciplina Cultura Organizacional

O usuário 3 realizou três sessões de busca, como pode-se verificar a seguir:

(Busca 1- Usuário 3)

<b>Procedimento para escolha da base de dados</b>	Endereço eletrônico (Portal CAPES - <a href="http://www.periodicos.capes.gov.br/">http://www.periodicos.capes.gov.br/</a> )		
<b>Base de dados</b>	Caderno de pesquisas em Administração		
<b>Reconhece facilmente o campo de busca</b>	SIM		
<b>Fonte dos descritores (termos de busca)</b>	Aulas em sala		
<b>Pontos de acesso utilizados</b>	Palavras-chave		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
1) CULTURA E TECNOLOGIA	simples	0	--
2) CULTURA	simples	0	--
3) CULTURA ORGANIZACIONAL	simples	0	--
<b>Reações e/ou emoções do usuário durante a busca</b>	1) “Demora em acessar a base; a maioria é em inglês, dificuldades; dificuldade em como interagir com o Portal, nem com a base selecionada; sai da base insatisfeito”.		
<b>Anotações do observador</b>	Tentando verificar como funciona a base; Estratégias feitas na hora; Saiu da base insatisfeito.		



(Busca 2 – Usuário 3)

<b>Procedimento para escolha da base de dados</b>	Endereço eletrônico ( <a href="http://www.rae.com.br/index.cfm">http://www.rae.com.br/index.cfm</a> )		
<b>Base de dados</b>	Revista de Administração de Empresas - <b>RAE</b>		
<b>Reconhece facilmente o campo de busca</b>	Não há campo de busca		
<b>Fonte dos descritores (termos de busca)</b>	Aulas em sala		
<b>Pontos de acesso utilizados</b>	Lista por Assunto		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
4) Não Há sistema de Busca	“Folheio de tela” (Screen browsing)	1	1
<b>Reações e/ou emoções do usuário durante a busca</b>	“demora para ver os artigos; não sabe o que fazer; não sabe buscar sobre isso”		
<b>Anotações do observador</b>	A base não tem campo de busca; tem que ser revista por revista; leu o resumo de 1 artigo, “linkou” o texto completo e salvou, mas é para outro trabalho.		

(Busca 3 – Usuário 3)

<b>Procedimento para escolha da base de dados</b>	Endereço eletrônico ( <a href="http://www.scielo.org/index.php?lang=en">http://www.scielo.org/index.php?lang=en</a> )		
<b>Base de dados</b>	SciELO		
<b>Reconhece facilmente o campo de busca</b>	SIM		
<b>Fonte dos descritores (termos de busca)</b>	Aulas em sala		
<b>Pontos de acesso utilizados</b>	Lista por assunto		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
5) CULTURA	simples	115	3
6) CULTURA AND TECNOLOGIA AND SIGNIFICADOS	avançada	0	0
<b>Reações e/ou emoções do usuário durante a busca</b>	“não se encontra nada”		
<b>Anotações do observador</b>	Após uma tentativa, desiste da busca		

**Tempo total de busca (usuário 3): 00: 50 min.**

**Caracterização e tema da busca do Usuário observado 4:**

<b>Pós-Graduação em</b>	Educação
<b>Formação acadêmica</b>	Comunicação
<b>Profissão</b>	Estudante
<b>Tema da busca</b>	Conceito de Excelência na Educação
<b>Objetivo da Busca</b>	Construção de Capítulo de Dissertação

O usuário 4 realizou duas sessões de busca, como pode-se verificar a seguir:

(Busca 1- Usuário 4)

<b>Procedimento para escolha da base de dados</b>	Ícone da área de trabalho		
<b>Base de dados</b>	Athena ( <i>Base Geral</i> ) – <a href="http://200.145.174.8:4505/ALEPH">http://200.145.174.8:4505/ALEPH</a>		
<b>Reconhece facilmente o campo de busca</b>	NÃO		
<b>Fonte dos descritores (termos de busca)</b>	Memória		
<b>Pontos de acesso utilizados</b>	Assunto		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
1) EXCELÊNCIA	simples	17	0
2) EXCELÊNCIA	Avançada (assunto)	0	0
<b>Reações e/ou emoções do usuário durante a busca</b>	“não se encontra nada”		
<b>Anotações do observador</b>	Após uma tentativa, desiste da busca		

(Busca 2 - Usuário 4)

<b>Procedimento para escolha da base de dados</b>	Ícone da área de trabalho			
<b>Base de dados</b>	Athena ( <i>Base local</i> ) – <a href="http://200.145.174.8:4505/ALEPH">http://200.145.174.8:4505/ALEPH</a>			
<b>Reconhece facilmente o campo de busca</b>	NÃO			
<b>Fonte dos descritores (termos de busca)</b>	Memória			
<b>Pontos de acesso utilizados</b>	Palavras-chave; Autor			
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>		<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
3) EXCELÊNCIA	simples		16	2
4) EXCELÊNCIAS	simples		1	0
5) MARIA ESTER FREITAS	Erro! de página	Avançada (autor)	0	0
6) QUALIDADE	simples		86	4
<b>Reações e/ou emoções do usuário durante a busca</b>	“Ícones, não encontra; resultados, não saber como proceder; 1º vez que acessa			
<b>Anotações do observador</b>	Não gosta buscar por artigos de periódicos; para isso prefere procurar o bibliotecário; apenas livros o usuário tenta procurar sozinho; mas até para anotar o nº de chamada na estante, tem dificuldades, não sabe onde clicar.			

**Tempo total de busca (usuário 4):** 00: 37 min.

**Caracterização e tema da busca do Usuário observado 5:**

<b>Pós-Graduação em</b>	Educação
<b>Formação acadêmica</b>	Administração de Empresas
<b>Profissão</b>	Professor
<b>Tema da busca</b>	Administração da Educação
<b>Objetivo da Busca</b>	Anti-projeto para Mestrado

(Busca1- Usuário 5)

<b>Procedimento para escolha da base de dados</b>	Ícone da área de trabalho		
<b>Base de dados</b>	Athena ( <i>Base local</i> ) – <a href="http://200.145.174.8:4505/ALEPH">http://200.145.174.8:4505/ALEPH</a>		
<b>Reconhece facilmente o campo de busca</b>	NÃO		
<b>Fonte dos descritores (termos de busca)</b>	Memória		
<b>Pontos de acesso utilizados</b>	Todos que a busca simples utiliza		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
TEORIA EDUCAÇÃO ADMINISTRAÇÃO	simples	5136	4
<b>Reações e/ou emoções do usuário durante a busca</b>	“tem problemas com a interface do sistema; não sabe localizar os itens”		
<b>Anotações do observador</b>	Grande dificuldade em lidar com o sistema; não sabe como obter o nº de chamada para localização do item na estante; busca muito imprecisa; muito “folheio” de telas.		

**Tempo total de busca** (usuário 5): 00: 22 min.

**Caracterização e tema da busca do Usuário observado 6:**

<b>Pós-Graduação em</b>	Educação
<b>Formação acadêmica</b>	pedagogia
<b>Profissão</b>	Estudante
<b>Tema da busca</b>	Comunicação alternativa para deficientes auditivos
<b>Objetivo da Busca</b>	Projeto de Mestrado

(Busca1- Usuário 6)

<b>Procedimento para escolha da base de dados</b>	Ícone da área de trabalho		
<b>Base de dados</b>	Athena ( <i>Base local</i> ) – <a href="http://200.145.174.8:4505/ALEPH">http://200.145.174.8:4505/ALEPH</a>		
<b>Reconhece facilmente o campo de busca</b>	SIM		
<b>Fonte dos descritores (termos de busca)</b>	Memória		
<b>Pontos de acesso utilizados</b>	Todos que a busca simples utiliza		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
1) COMUNICAÇÃO ALTERNATIVA DEFICIENTES FÍSICOS	simples	774	2
2) COMUNICAÇÃO ALTERNATIVA DEFICIENTES FÍSICOS SURDOS	simples	792	6
<b>Reações e/ou emoções do usuário durante a busca</b>	“não tem paciência com o sistema”		
<b>Anotações do observador</b>	Estratégias de busca feitas na hora; refinar busca		

**Tempo total de busca** (usuário 6): 00: 21 min.

**Caracterização e tema da busca do Usuário observado 7:**

<b>Pós-Graduação em</b>	Educação
<b>Formação acadêmica</b>	pedagogia
<b>Profissão</b>	Bolsista
<b>Tema da busca</b>	História da leitura escolar no Brasil
<b>Objetivo da Busca</b>	Para a dissertação

(Busca1- Usuário 7)

<b>Procedimento para escolha da base de dados</b>	Ícone da área de trabalho		
<b>Base de dados</b>	Athena ( <i>Base local</i> ) – <a href="http://200.145.174.8:4505/ALEPH">http://200.145.174.8:4505/ALEPH</a>		
<b>Reconhece facilmente o campo de busca</b>	SIM		
<b>Fonte dos descritores (termos de busca)</b>	Memória		
<b>Pontos de acesso utilizados</b>	Assunto		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
1) LEITURA ESCOLAR HISTÓRIA BRASIL	avançada (assunto)	1	1
2) ARROYO	avançada (autor)	03	1
3) LAJOLO	avançada (autor)	12	5
<b>Reações e/ou emoções do usuário durante a busca</b>	“prefere os livros”		
<b>Anotações do observador</b>	Prefere ir pelo autor.		

**Tempo total de busca (usuário 7):** 00: 15 min.

**Caracterização e tema da busca do Usuário observado 8:**

<b>Pós-Graduação em</b>	Filosofia
<b>Formação acadêmica</b>	Filosofia
<b>Profissão</b>	Estudante
<b>Tema da busca</b>	Auto organização
<b>Objetivo da Busca</b>	Para a dissertação

(Busca1- Usuário 8)

<b>Procedimento para escolha da base de dados</b>	Ícone da área de trabalho		
<b>Base de dados</b>	Athena ( <i>Base local</i> ) – <a href="http://200.145.174.8:4505/ALEPH">http://200.145.174.8:4505/ALEPH</a>		
<b>Reconhece facilmente o campo de busca</b>	SIM		
<b>Fonte dos descritores (termos de busca)</b>	Memória		
<b>Pontos de acesso utilizados</b>	1º Todos que a busca simples utiliza; 2º Título		
<b>Combinações de termos (descritores) de busca e operadores (AND; OR; NOT; etc)</b>	<b>Tipo de Busca</b>	<b>Registros recuperados</b>	<b>Referências relevantes recuperadas</b>
1) AUTO ORGANIZAÇÃO SUJEITO	simples	483	0
2) ORGANIZAÇÃO SUJEITO	Avançada (Título)	0	0
3) AUTO ORGANIZAÇÃO	Avançada (Título)	0	0
4) AUTO-ORGANIZAÇÃO	simples	13	3
<b>Reações e/ou emoções do usuário durante a busca</b>	“consultar novamente o orientador”		
<b>Anotações do observador</b>	Indeciso quanto aos termos de busca; prefere consultar os livros do acervo local		

**Tempo total de busca (usuário 8): 00: 25 min.**

De acordo com os dados obtidos, ao menos com a observação destes usuários, verificou-se primeiramente que em média o tempo de busca dos usuários ficou em torno de 00:28,5 minutos, sendo que o usuário 7 gastou 00:15 minutos e o usuário 3 gastou 00:50 minutos para efetuar a sua busca.

A escolha da base de dados ficou em detrimento apenas à base de dados catalográfica **Athena**, pois 5 usuários consultaram somente a base de dados do acervo local, 1 usuário consultou a base Athena local e a geral, e 1 usuário consultou o acervo local e a Internet através da ferramenta de busca **Google**. Somente 1 usuário optou por consultar as bases de dados do Portal de periódicos **CAPES** e da Revista de Administração de Empresas – **RAE**. Vê-se aqui um dos principais problemas apontados pelos próprios usuários que responderam ao questionário que é a escolha da base de dados (ver Tabela 4) e o principal motivo para procurarem o bibliotecário (ver Tabela 5). Segundo Cuenca (1999, p. 292),

Além do desconhecimento, dificuldades como: a existência de várias interfaces de busca para o acesso às bases de dados, tempo de busca, campos disponíveis para recuperação e a não-familiarização com o vocabulário especializado da área; são colocados como motivo da não-utilização do acesso às bases automatizadas pelos usuários de buscas informatizadas.

Quanto a isso, o bibliotecário de referência deve-se dispor de mecanismos que facilitem a divulgação das bases de dados oferecidas, ou seja, planejar “[...] novos instrumentos para administrar e acessar a informação” (FIGUEIREDO, 1999, p. 94), despertando a importância que as mesmas tem para o desenvolvimento das pesquisas, pois conforme apontado por 4 dos usuários observados, preferem obter informações somente através de livros em detrimento da vasta gama de periódicos e artigos de periódicos



disponíveis em diversas bases de dados atualmente oferecidas que poderiam enriquecer ainda mais as pesquisas, principalmente pela dinâmica que os artigos de periódicos possuem.

Hernández Salazar (2003, p. 172) afirma ainda que “[...] as comunidades que utilizam em menor medida recursos tecnológicos de informação para realizar suas atividades são as que estudam fenômenos circunscritos na área de humanidades”<sup>51</sup>.

Neste sentido é necessário que se considere o processo cognitivo relacionado com o acesso e uso da informação dos indivíduos que se deseja capacitar, o que implica identificar suas características, ou seja, seus perfis de necessidades de informação e a formação no uso dos recursos tecnológicos, além do comportamento que seguem quando buscam informação (HERNÁNDEZ SALAZAR, 2003).

No que concerne ao comportamento de busca Hernández Salazar (2003), diz que os usuários advindos da área de humanas se caracterizam pela pouca utilização de ferramentas secundárias e conseqüentemente tecnológicas e quando o fazem, as mais utilizadas são resenhas, bibliografias e catálogos. Costumam recorrer a vários meios informais de comunicação como consulta ou troca de idéias com colegas, buscas em suas coleções pessoais e recorrem a sua memória. Nesta pesquisa, 6 dos 8 alunos observados, recorreram a memória quanto à fonte dos descritores (termos de busca) para a aplicação nas estratégias de busca. Os sistemas consultados não disponibilizavam nenhum tipo de Tesouro.

Em relação ao reconhecimento do(s) campo(s) de busca, 6 usuários visualizaram facilmente, enquanto que 2 tiveram dificuldades em localizar o campo de busca. Já em relação às estratégias de busca utilizadas, das 32 estratégias de busca formuladas (combinação de termos de busca), apenas 3 (9,4%) do total utilizou-se de operadores booleanos, mas somente do conector ‘E’, ou seja, os dois usuários que fizeram buscas em outros sistemas, visto que a

---

<sup>51</sup> [...] las comunidades que utilizan en menor medida recursos tecnológicos de información para realizar sus actividades son las que estudian fenómenos circunscritos al área de humanidades.

base de dados Athena não comporta tais operadores (E, OU, NÃO). Já o tipo de busca utilizada (simples ou avançada), das 12 buscas realizadas pelos 8 usuários, 21 combinações de termos de busca (63,6%) foram feitas em buscas simples, enquanto que 12 combinações de termos de busca (36,4%) foram feitas utilizando os recursos de busca avançada como autor, título, assunto e preferências por idioma e formato do arquivo (no caso do Google). Uma das bases consultadas não havia sistema de busca (RAE), o usuário “folheou” as telas.

Pode-se perceber a sub-utilização dos recursos disponíveis. De acordo com Tasso et al. (2002), alguns usuários não exploram, toda a potencialidade do sistema de recuperação da informação, mesmo se a busca progride, e muitos usuários adotam uma estratégia de busca única para uma sessão inteira, com conseqüências negativas e quando adotam uma estratégia válida, mas não familiar, erros secundários (por exemplo erros de digitação, ou de combinação incorreta dos termos) podem causar o seu abandono.

No que tange a relação registros visualizados e registros relevantes recuperados (no final da busca executada), pode-se ver isso na tabela que se segue:

**Tabela 9 - nº de registros relevantes recuperados por nº de registros recuperados**

Usuário	Busca	nº de registros relevantes recuperados	nº de registros recuperados	Coefficiente de Precisão
<b>1</b>	1	26	28282	0,0009
	2	16	143	0,14
<b>2</b>	1	3	9477	0,0003
<b>3</b>	1	0	0	0
	2	1	1	1
	3	3	115	0,026
<b>4</b>	1	0	17	0
	2	6	103	0,06
<b>5</b>	1	4	5136	0,0008
<b>6</b>	1	8	1566	0,0051
<b>7</b>	1	7	16	0,43
<b>8</b>	1	3	496	0,006

Onde o coeficiente de precisão é dado por:

$$P = \frac{drr}{drr + dnrr} \quad \text{onde:}$$

$P \rightarrow$  coeficiente de precisão;

$drr \rightarrow$  documentos relevantes (pertinentes) recuperados;

$dnrr \rightarrow$  documentos não relevantes recuperados.

De acordo com a tabela acima, o usuário 3 em sua 2ª busca obteve 100% de aproveitamento em sua busca, entretanto foi apenas 1 item visualizado e recuperado. O usuário 7 obteve maior precisão em suas buscas com o coeficiente de precisão de **0,43**, enquanto que o usuário 2 obteve a menor precisão (coeficiente de **0,0003**) entre os observados durante o processo de busca.

Ser capaz de encontrar qualquer documento relevante é ser capaz de articular a necessidade de informação nos termos usados nos documentos. Mas em nenhum momento os usuários aproveitaram os documentos recuperados para refinar a busca. Conforme Tasso et al. (2002), os usuários, depois de recuperar os documentos, eles não olham atentamente para os mesmos, perdendo a oportunidade de melhorar o processo de busca.

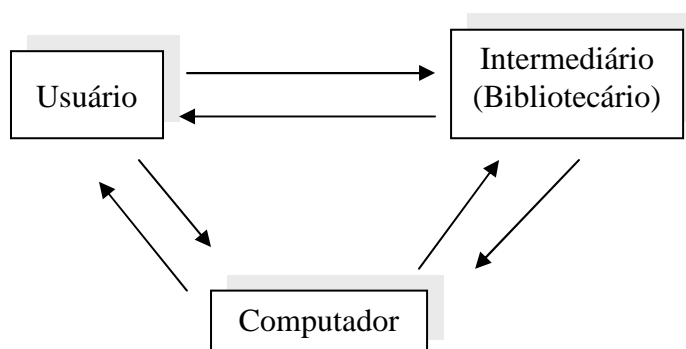
Até mesmo quando os usuários lêem os documentos recuperados, eles não são capazes de explorar a informação que já adquiriram: por exemplo, eles julgam um documento como relevante sem perceber que poderiam extrair alguns termos para melhor reformular suas questões (TASSO et al., 2002, p. 345)<sup>52</sup>.

---

<sup>52</sup> Even when users read the retrieved document, they are not always able to exploit the information they gathered: for instance, they judge a document as relevant without realizing that they could extract some terms to better reformulate their query.

Neste sentido, “é evidente que a habilidade dos usuários para articular conceptualmente necessidades de informação e expressá-las com questões [de busca] afetam o sucesso da busca” (PENNANEN; VAKKARI, 2003, p. 759)<sup>53</sup>.

Outro ponto importante é que em nenhum momento, os usuários recorreram ao bibliotecário. Segundo Hernández Salazar (2003, p. 173) os usuários “[...] recorrem ao bibliotecário como última possibilidade para encontrar informação”<sup>54</sup>. O que não deveria acontecer já que o intermediário bibliotecário auxilia seus usuários a definir o problema de busca, a escolher a melhor base de dados que poderá responder a questão, ajuda a definir os termos de busca, a melhor estratégia a ser adotada, ou seja, interagindo com as bases de dados e o usuário para melhor atender as necessidades de informação deste último, conforme a figura que se segue.



**Figura 20:** Modelo de interação na recuperação da informação (adaptado)

**Fonte:** Foster et al., 2002, p. 884

Portanto, uma maior divulgação dos serviços de referência a comunidade usuária poderia proporcionar a conscientização das contribuições que o profissional bibliotecário pode oferecer na utilização dos recursos informacionais disponíveis.

A seguir serão apresentadas as considerações finais da pesquisa.

<sup>53</sup> It is evident that users' ability to articulate information needs conceptually and Express them as queries affects the search success.

<sup>54</sup> [...] recurren al bibliotecario como última posibilidad para encontrar información.

## CAPÍTULO 7

A questão de por que existe algo ao invés de nada deve  
sempre inspirar nossa humildade.  
GLEISER<sup>55</sup>

### CONSIDERAÇÕES FINAIS

A automatização das bibliotecas não foi assim tão drástica de se realizar, embora árduo, uma vez que não desafiou ou ameaçou o papel das bibliotecas e dos bibliotecários, e sim dinamizou a prestação de serviços e produtos informacionais.

Conforme Rodríguez Bravo e Santos da Paz (2002), o que caracterizou as bibliotecas nesta fase foi a automatização de alguns processos, fundamentalmente a Catalogação.

Porém, como o mundo da informação em geral vem se tornando cada vez mais digital, interconectado e dominado pela Web, todos os profissionais e setores que lidam com a informação devem prestar atenção em como a sempre ascendente conectividade e a digitalização de recursos estão afetando seus trabalhos, sua profissão, e a comunidade a qual eles servem (JANES, 2002).

Neste sentido, segundo Figueiredo (1999, p. 91),

Os papéis dos bibliotecários de referência mudaram um pouco, mas eles ainda continuam como bibliotecários de referência. A transição que se tem pela frente, agora, da biblioteca automatizada para a eletrônica não será [e não esta sendo] tão simples e fácil. É uma mudança muito mais

---

<sup>55</sup> Gleiser, Marcelo. **A dança do universo**: dos mitos de criação ao Big Bang. São Paulo: Companhia das Letras, 1997.

fundamental, não somente da natureza do serviço de referência, mas também da natureza da própria biblioteca.

Isto é, hoje a biblioteca digital é pensada como uma nova estratégia para o resgate de informações onde o texto completo de documentos está disponível on line (OHIRA; PRADO, 2002).

Mediante isto os clientes/usuários têm conseguido suprir suas necessidades informacionais eficientemente, sem a intermediação do profissional bibliotecário, neste emergente meio digital de bases e bancos de dados online ?

Ao longo desta pesquisa verificou-se que diversos problemas impedem que os clientes/usuários utilizem-se de toda a potencialidade que as bases de dados atualmente oferecem para a recuperação da informação. Conforme pode-se constatar, principalmente a escolha de bases de dados; a seleção dos termos; a elaboração e a aplicação das estratégias de busca; a utilização dos operadores lógicos; a utilização dos Tesouros, etc; além dos próprios problemas de interação do usuário com o bibliotecário e as interfaces dos sistemas de recuperação da informação e do desconhecimento dos benefícios que tais ferramentas podem oferecer para a realização de suas atividades acadêmicas, foram apontadas como as principais dificuldades.

Nas comunidades das áreas de Humanas, principalmente (ou o que se pode perceber neste estudo), realmente existe o pouco preparo ou mesmo a falta de interesse dos usuários para utilizarem-se destes recursos, “[...] sobre tudo os que se encontram na Internet ou na WWW”<sup>56</sup> (HERNÁNDEZ SALAZAR, 2003, p. 173). Tais comunidades não percebem os benefícios adicionais que os sistemas de recuperação da informação atuais oferecem, não os explorando em sua máxima potencialidade (HERNÁNDEZ SALAZAR, 2003).

---

<sup>56</sup> [...] sobre todo los que se encuentran en Internet o la WWW.

Muito se tem discutido sobre a capacitação e/ou instrução dos usuários para uso dos recursos tecnológicos de informação atualmente disponíveis. Conforme Hernández Salazar (2003, p. 173) os usuários,

Requerem ser formados nas ferramentas para buscar e recuperar informação, catálogos online, bases dados e sistemas de informação arquitetados na Internet e na World Wide Web (WWW), como motores de busca, listas de discussão, acesso e transferência de arquivos remotos<sup>57</sup>.

Entretanto, aos usuários das áreas de Humanas, conforme afirma Hernández Salazar (2003, p. 174),

Não lhes interessa seguir um processo de aprendizagem formal, porque não consideram dedicar tempo a isto e mais ainda não tem a disposição para aprender sobre os recursos tecnológicos, os membros dessas comunidades esperam ser capazes de usá-los sem ajuda ou apoio diretos de um bibliotecário ou profissional da informação<sup>58</sup>.

Neste sentido, Figueiredo (1999, p. 92) destaca que,

---

<sup>57</sup> Requieren ser formados en las herramientas para buscar y recuperar información, catálogos en línea, bases de datos y sistemas de información montados en Internet y la World Wide Web (WWW), como motores de búsqueda, listas de discusión, acceso u transferencia de archivos remotos.

<sup>58</sup> No les interesa seguir un proceso de aprendizaje formal, por lo que no han considerado dedicar tiempo a esto y más aún no tienen la disposición para aprender sobre los recursos tecnológicos, los miembros de estas comunidades esperam ser capaces de usarlos sin ayuda o apoyo directos de un bibliotecario o profesional de la información.

Ensinar como fazer uso dos recursos da biblioteca um-a-um tornou-se uma atividade básica na referência, mas essa abordagem já atingiu o limite e não é mais uma estratégia a ser mantida. A instrução no uso da biblioteca eletrônica deve se tornar uma parte central do núcleo do currículo da universidade; deve ser, portanto um ensino regular ministrado pelo corpo docente, e não mais constituir uma carga extra para o bibliotecário de referência. Embora seja aconselhável que muito do corpo docente seja bibliotecário.

Efetivamente, a biblioteca digital constitui um dispositivo fundamental entorno do qual está se reelaborando a teoria biblioteconômica para a sociedade da informação (GARCÍA MARCO, 2002). Para Agustín Lacruz (1998), a biblioteca digital é nada mais nada menos que a resposta da interação que se estabelece entre o ambiente da instituição documental, conhecida como biblioteca, com as necessidades de seus usuários, no contexto da sociedade da informação.

É neste ambiente que, o desenvolvimento de instrumentos para administrar, acessar, recuperar a informação e que possam ser utilizados com o mínimo de intermediação, é competência também dos bibliotecários de referência, ou como diz Figueiredo (1999, p. 91) “o que os bibliotecários de referência deviam estar fazendo é o desenho de novos sistemas; esses bibliotecários são planejadores naturais de novas ferramentas para a administração, o acesso, a recuperação e a distribuição da informação”, pois,

Se a informação for disposta de maneira organizada, será fácil para o usuário recuperar o que estiver procurando. Por outro lado, se a estrutura apresentar uma trilha de navegação desorganizada ou um design



cognitivamente pobre, o folheio (*browsing*) ou a pesquisa mais direccionada por parte do usuário se torna difícil ou ineficiente (LIMA, 2004, p. 126).

Sendo assim, ao bibliotecário cabe dominar as ferramentas tecnológicas de informação e comunicação para melhor prover seus usuários de informações pertinentes, pois

Os bibliotecários de referência são o único elo com os usuários; aprenderam com o sistema on-line, sabem o que funciona e não funciona e têm idéia de como pode funcionar melhor; eles sabem que os usuários não têm mais tempo do que eles próprios para dominar ferramentas de pesquisa, porque a tecnologia da informação é um negócio dos bibliotecários, não dos usuários (FIGUEIREDO, 1999, p. 91).

A necessidade de desenvolvimento de interfaces de busca (embora muito já tenha sido feito neste sentido) mais eficientes é evidente, devido à gama crescente de informação disponível em meio digital. Os sistemas de recuperação da informação deveriam ser desenhados para o usuário extrair informação da base de dados, termos para refinar a busca, e dar assistência nos julgamentos de relevância (FOSTER et al., 2002), pois,

Do ponto de vista da interação homem-computador, a navegação é o resultado da interação entre os elementos do sistema e as necessidades dos usuários. Esta interação, que ocorre através de uma interface entre o sistema e o usuário, pode ser física, perceptiva e cognitiva. Assim, a navegação no hipertexto é uma ação que pressupõe aspectos mecânicos, cognitivos e tecnológicos em um só processo (LIMA, 2004, p. 127).

Neste sentido, os estudos sobre o comportamento de busca dos usuários finais e as interações destes com os sistemas de recuperação da informação são primordiais para o desenvolvimento de ferramentas de busca mais eficientes. Ou seja, pesquisas que abordem o comportamento de busca e a interação dos usuários com os sistemas de recuperação da informação, são necessárias principalmente por duas razões: para que se assimile mais o que está acontecendo de fato na prática, e tentar usar o que for encontrado nessas pesquisas para o design e desenvolvimento de melhores processos e interfaces de recuperação da informação (SPINK; SARACEVIC, 1997).

Somente assim, acredita-se, que os clientes/usuários, dominando os processos de planejamento e operacionalização das estratégias de busca e interagindo com interfaces de buscas mais eficientes, serão capazes de recuperar a informação necessária com maior efetividade.

Com esta pesquisa, objetivou-se identificar o perfil e as necessidades informacionais dos usuários, verificando como os recursos disponíveis estão sendo utilizados pelos usuários além de, identificar as necessidades de otimização dos processos de planejamento e operacionalização das estratégias de busca e das interações entre bibliotecário, usuário e base de dados, para uma efetiva recuperação da informação. Pois, é a partir desses dados, que o bibliotecário poderá traçar o planejamento adequado, utilizando-se das tecnologias de informação e comunicação atualmente empregadas, para o desenvolvimento e administração de serviços e produtos informacionais mais eficientes.

Espera-se que esta pesquisa venha a trazer subsídios a área da Biblioteconomia e da Ciência da Informação, contribuindo no que se refere à busca e recuperação da informação por usuários finais em bases de dados especializadas, suscitando outras discussões sobre o assunto para que, desta forma, se de continuidade a este tema atual e imprescindível para a área.

## REFERÊNCIAS.

- AGUSTÍN LACRUZ, María del Carmen. Bibliotecas digitales y sociedade de la información. **Scire: representación y organización del conocimiento**. Universidad Zaragoza, v.4, n.2, p. 47-62, jul./dic. 1998.
- ANGÓS ULLATE, José Maria; SALVADOR OLIVÁN, José Antonio. Metadatos em los documentos HTML. **Scire: representación y organización del conocimiento**. Universidad Zaragoza, v.5, n.2, p. 47-61, jul./dic. 1999.
- ANGÓS ULLATE, José Maria; SALVADOR OLIVÁN, José Antonio; FERNÁNDEZ RUÍZ, Maria Jesús. Evolución em la recuperación de información: de la línea de comandos a la interfaz gráfica Web. **Scire: representación y organización del conocimiento**. Universidad Zaragoza, v.7, n.1, p. 127-142, ene./jun. 2001.
- ARELLANO, Miguel Ángel Márdero. Serviços de referência Virtual. **Ciência da informação**, Brasília, v.30, n.2, p.7-15, maio/ago. 2001.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR 12676**: métodos para análise de documentos: determinação de seus assuntos e seleção de termos de indexação. Rio de Janeiro, ago. 1992.
- BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern Information Retrieval**. New York: ACM Press, 1999. 513p.
- BALBY, C. N. Formatos de intercâmbio de registros bibliográficos: conceitos básicos. **Cadernos da F.F.C. Marília**, v.4, n.1, 1995. p. 29-35.
- BAINS, Simon. End-user searching behavior: considering methodologies. **The Katharine Sharp Review**. n. 4, 1997.
- BATISTA. **Armazenamento e recuperação de informações textuais**. Disponível em: <<http://www.dsc.ufcg.edu.br/~baptista/cursos/BDMM/Capitulo3A.doc>>. Acesso em: 25 out. 2004.
- BERTHOLINO, Maria Luzia Fernandes. Buscas em bases de dados. In: RAMOS, Maria Etelvina Madalozzo (org.). **Tecnologia e novas formas de gestão em bibliotecas universitárias**. Ponta Grossa : UEPG, 1999. p. 145-155.
- BRAGA, Gilda Maria. Informação, ciência da informação: breves reflexões em três tempos. **Ciência da Informação**, Brasília, v.24, n.1. 1995.
- CARDOSO, Olinda Nogueira Paes. **Recuperação de Informação**. Disponível em: <<http://www.dcc.ufpa.br/infocomp/artigos/v2.1/olinda.pdf>>. Acesso em: 21 nov. 2004.
- CAMARGO R. **Interatividade através de homepage de bibliotecas na era digital**. Disponível em: <<http://libdigi.unicamp.br/document/?down=8303>>. Acesso em: 21 nov.2004.

CAREGNATO, Sônia Eliza. Busca e uso de informação por alunos de pós-graduação: implicações para o desenvolvimento de habilidades informacionais na área de comunicação. In: Sociedade brasileira de estudos interdisciplinares da comunicação, Congresso brasileiro de ciências da comunicação, 26., 2003, Belo Horizonte. **Anais**. Belo Horizonte, 2003. 13 p.

CARO-CASTRO, Carmen; CEDEIRA SERANTES, Lucía; TRAVIESO RODRÍGUES, Crispulo. La investigación sobre recuperación de información desde la perspectiva centrada en el usuario: métodos y variables. **Revista Española de Documentación Científica**. Madrid, v.26, n.1, p. 40-55, ene./mar. 2003.

CASTELLS, Manuel. **A sociedade em rede**. 2.ed. Tradução: Roneide Venâncio Majer. São Paulo: Paz e Terra S.A., v. 1, 1999. (Tradução do original: **The rise of the network society**).

CHOWDHURY, Gobinda G. **Digital libraries and reference services: present and future**. Journal of Documentation, v.58, n.3, p.258-283. 2002.

CIANCONI, R. B. Banco de dados de acesso público. **Ciência da Informação**. Brasília, v.16, n.1, p. 53-59, jan./jun. 1987.

CINTRA, Ana Maria et al. **Para entender as linguagens documentárias**. 2.ed. rev. ampl. São Paulo: Polis, 2002. 96p.

COHEN, Max F. Alguns aspectos do uso da informação na economia da informação. **Ciência da informação**, Brasília, v.31, n.3, p.26-36, set./dez. 2002.

CORREA DA CRUZ, Angelo Antonio Alves et al. Impacto dos periódicos eletrônicos em bibliotecas universitárias. **Ciência da Informação**. Brasília, v.32, n.2, p. 47-53, maio/ago. 2003.

CUENCA, Angela Maria Belloni. O usuário final da busca informatizada: avaliação da capacitação no acesso a bases de dados em biblioteca acadêmica. **Ciência da Informação**. Brasília, v.28, n.3, p. 291-299, set./dez. 1999.

DEBOWSKI, Shelda. Wrong way: go back! An exploration of novice search behaviours while conducting an information search. **The Electronic Library**. v.19, n.6, p.371-382, 2001.

DIAS, Ataíde Guilherme. Periódicos eletrônicos: considerações relativas à aceitação deste recurso pelos usuários. **Ciência da Informação**. Brasília, v.31, n.3, p. 18-25, set./dez. 2002.

DIAS, Claudia. **Usabilidade na web**: criando portais mais acessíveis. Rio de Janeiro: Alta Books, 2003.

DICIONÁRIO AURÉLIO ELETRÔNICO: século XXI. Rio de Janeiro: Nova Fronteira, 2002.

FERNEDA, Edberto. **Recuperação de Informação**: análise sobre a contribuição da Ciência da Computação para a Ciência da Informação. Tese (Doutorado) – Escola de Comunicação e Artes. USP, São Paulo, 2003.

FIQUEIREDO, Nice Menezes de. **Paradigmas modernos da Ciência da informação**. São Paulo: Polis, 1999, 168 p.

FLAMINO, Adriana Nascimento. **MARC21 e XML como ferramentas para a consolidação da catalogação cooperativa automatizada** : uma revisão de literatura. 2003. 142f. TCC (trabalho de conclusão de curso) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2003.

FLAMINO, Adriana Nascimento; SANTOS, Plácida L. V. A. Costa. MARC21 e XML como ferramentas para a consolidação da catalogação cooperativa automatizada: uma revisão de literatura. In: VIDOTTI, Ap. B. Gregório (coord). **Tecnologia e conteúdos informacionais: abordagens teóricas e práticas**. São Paulo: Polis, 2004. p. 114-138.

FOSTER, Allen; SPINK, Amanda; ELLIS, David; LAM, H. M.; FORD, Nigel; BURTON, R.; WILSON, T. D. Information seeking and mediated searching. Part 5. user-intermediary interaction. **Journal of the American Society for Information Science and Technology**. v. 53, n. 11, p. 883-893, 2002.

FRANTZ, V.; SHAPIRO, J.; VOISKUNSKII, V. **Automated Information Retrieval: Theory and Methods**. San Diego, CA: Academic Press, 1997. 365 p.

FUJITA, Mariângela Spotti Lopes. A identificação de conceitos no processo de análise de assunto para a indexação. **Revista Digital de Biblioteconomia e Ciência da Informação**. Campinas, v.1, n.1, p. 60-90, jul./dez. 2003.

FURGERI, S. **Utilização de catálogos XML para o desenvolvimento do comércio eletrônico**. 1999. 92 f. Dissertação (Mestrado em Sistemas de Informação) - Pontifícia Universidade Católica de Campinas, Instituto de Informática, Campinas, 1999.

FURRIE, B. **O MARC bibliográfico**: um guia introdutório; catalogação legível por computador. Brasília : Thesaurus, 2000. 95 p. Tradução de: Beatriz Valadares Cendón, Sonia Burnier, Maria Helena Santos e Natália Guiné de Mello Carvalho.

GARCÍA MARCO, Francisco Javier. La vitalidad de las ciencias de la documentación. **Scire: representación y organización del conocimiento**. Universidad Zaragoza, v.8, n.2, p. 9-21, jul./dic. 2002.

GARCÍA MARTÍNEZ, Ana M<sup>a</sup>. Definición y estilo de los objetos de información digitales y metadatos para la descripción. **Boletín de la Asociación Andaluza de Bibliotecarios**. [S. l.], n. 63, p. 23-47, Junio. 2001.

GONÇALVES, G. J. **Sistemas de informação automatizados**: uma análise crítica sobre a sua eficácia. Dissertação (mestrado) – Instituto de Administração e Gerência. PUC, Rio de Janeiro, 1996.

GONZÁLEZ OLIVARES, José Luis; GARCÍA GÓMEZ, Juan Carlos. Aproximación a la evaluación cuantitativa de los sistemas de recuperación de información de la prensa en Internet: Exhaustividad y precisión. **Scire: representación y organización del conocimiento**. Universidad Zaragoza, v.7, n.1, p. 143-152, ene./jun. 2001.

GRIGÓRIO, Deise de Araújo. **Teses e Dissertações em saúde pública e portarias do Ministério da Saúde: estudos de uma década sob a ótica do Decs.** Dissertação (mestrado) – Escola Nacional de Saúde Pública. Fundação Oswaldo Cruz, Rio de Janeiro, ago. 2002.

GROGAN, Denis. **A prática do serviço de referência.** Tradução: Antonio Agenor Brinquet de Lemos, Brasília: Brinquet de Lemos/Livros, 1995.

GUINCHAT, Claire; MENOU, Michel. **Introdução geral às ciências e técnicas da informação e documentação.** Tradução: Miriam Vieira da Cunha. 2.ed. corr. aum. Brasília: IBICT, 1994. 540p.

GÜNTHER, Hartmut. Como elaborar um questionário. In: PASQUALI, Luiz. **Instrumentos psicológicos: manual prático de elaboração.** Brasília: LABPAM; IBAPP, 1999. p. 231-258.

HAVERKAMP, D.S., GAUCH S. Intelligent Information Agents: Review and Challenges for Distributed Information Sources. **Journal Of The American Society for Information Science**, v. 49, n.4, p. 304-311, 1998.

HEEMANN, V. **Avaliação ergonômica de interfaces de bases de dados por meio de Checklist especializado.** 1997. Dissertação (Mestrado em Engenharia da Produção) – Universidade Federal de Santa Catarina, Florianópolis, 1997. Disponível em: <<http://www.eps.ufsc.br/disserta97/heemann/>>. Acesso em: 25 Set. 2003.

HERNÁNDEZ SALAZAR, Patricia. La formación de usuarios de tecnologías de información: um prototipo para humanistas. **Scire: representación y organización del conocimiento.** Universidad Zaragoza, v.9, n.2, p. 169-186, ene./jun. 2003.

INGWERSEN, P. **Information retrieval interaction.** London: Taylor Graham, 1992.

JAMIL, George Leal; NEVES, Jorge Tadeu de Ramos. A era da informação: considerações sobre o desenvolvimento das tecnologias da informação. **Perspectivas em ciência da informação**, Belo Horizonte, v.5, n.1, p. 41-53, jan./jun. 2000.

JANES, Joseph. Digital references: reference librarians' experiences and attitudes. **Journal of the American Society for Information Science and Technology**. v.53, n.7, p. 549-566, may. 2002.

JANTZ, Ronald C. Innovaciones tecnológicas en la biblioteca: proyectos digitales que ofrecen nuevas oportunidades al bibliotecario y a la biblioteca. **Boletín de la Asociación Andaluza de Bibliotecarios**, n. 67, p. 61-69, jun. 2002.

KOWALSKI, G. **Information Retrieval Systems: theory and Implementation.** Kluwer Academic Publishers, 1997. 282 p.

LANCASTER, F.W. **Indexação e resumos: teoria e prática.** Brasília: Brinquet de Lemos, 1993. 347 p.

LARGE, Andrew; MOUKDAD, Haidar. Users perceptions of the Web as revealed by transaction log analysis. **Online Information Review**. v. 25, n.6, p. 349-358, 2001.

LE COADIC, Yves-françois. **A ciência da informação**. Brasília: Briquet de Lemos Livros, 1996. (Tradução do Original francês **La science de l'information**. Paris: PUF, 1994 – Collection que sais-je?).

LEITÃO, S. P; ROSSI, L. P. Natureza da mudança nas organizações: uma introdução crítica. **Revista Arché Interdisciplinar**, v.10, n.28, p.13-38, 2000.

LEVACOV, Marília. Bibliotecas virtuais: (r)evolução?. **Ciência da informação**, Brasília, v. 26, n. 2, maio/ago. 1997.

LEWIS, D. D.; JONES, K. S. Natural Language Processing for Information Retrieval. **Communications of the ACM** , v.39, n.1, p.92-101. 1996

LIMA, Gercina Ângela Borém. A navegação em sistemas de hipertexto e seus aspectos cognitivos. **Cadernos BAD**, n.1, p. 126-138, 2004.

LOPES, Ilza Leite. Uso das linguagens controlada e natural em bases de dados: revisão de literatura. **Ciência da informação**, Brasília, v.31, n.1, p.41-52, jan./abr. 2002.

LOPES, Ilza Leite. Estratégia de busca na recuperação da informação: revisão da literatura. **Ciência da informação**, Brasília, v.31, n.2, p.60-71, maio/ago. 2002.

LOPEZ ALONSO, Miguel-Ángel . Relevancia de las recuperaciones con texto completo. **Organización del conocimiento en sistemas de información y documentación**, n.3, p. 275-293, 1999.

MA, Wei. A database selection expert system based on reference librarian's database selection strategy: a usability empirical evaluation. **Journal of the American Society for Information Science and Technology**. v.53, n.7, p. 567-580, may. 2002.

MACIEL, J. **Elementos da teoria geral dos sistemas**. Petrópolis: Vozes, 1974. 404 p.

MANUEL AGUADO, Pedro. Los sistemas expertos y la recuperación documental: ejemplos de aplicación. **Scire: representación y organización del conocimiento**. Universidad Zaragoza, v.1, n.2, p. 21-32, jul./dic. 1995.

**MARC21**: formato condensado para dados bibliográficos. Tradução e adaptação: Margarida M. Ferreira. 2. ed. Marília: Unesp – Marília – Publicações, 2002. v.1. 247 p. Tradução e adaptação de: MARC 21 concise for bibliographic data da Network development and MARC Standards Office – Library of Congress – USA.

MARCHIORI, Patrícia Zeni. Para além das bibliotecas: o bibliotecário como information broker. **Associação Brasileira de ensino de Biblioteconomia e Documentação**. 1998. Documentos ABEBD, 9.

MARTÍNEZ GONZÁLEZ, M. M. Extended Markup Language (XML) : Una solución para modelar documentos y sus interrelaciones basada en la semántica de la información. **Scire: Representación y Organización del Conocimiento**, Zaragoza, v.6, n.2, p.121-151, jul./dic. 2000.

MARTUCCI, Elizabeth Márcia. Revisitando o Trabalho de referência: uma contribuição teórica para a abordagem interpretativa de pesquisa. **Perspectivas em ciência da informação**. Belo Horizonte, v.5, n.1, p. 99-115, jan./jun. 2000.

MCCHOMBU, Kingo; MENOU, Michel. Os profissionais da informação em comunidades desfavorecidas. In: VALENTIM, Marta Lúcia (Org.). **Atuação profissional na área de informação**. São Paulo, Polis, 2004, p. 129-150

MEADOWS, A. J. **A comunicação científica**. Tradução: Antonio Agenor Briquet de Lemos. Brasília: Briquet de Lemos/ Livros, 1999.

MENDES, Gildásio, TACHIZAWA, Takeshy. **Como fazer monografia na prática**. 6.ed. rev. ampl. Rio de Janeiro: Editora FGV, 2001.

MERCADO, Heidi. Library instruction and online database searching. **Reference Services Review**. v. 27, n. 3, p. 259-265, 1999.

MILANESI, Luís. A formação do informador. **Informação & Informação**. Londrina, v.7, n.1, p.07-40, jan./jun. 2002.

MILSTEAD, J. ; FELDMAN, S. Metadata: Cataloging by Any Other Name. **Online**, Jan. 1999. Disponível em: <<http://www.onlineinc.com/online/OL1999/milstead1.html>>. Acesso em: 10 set. 2003.

MODESTO, Fernando; MACEDO, Neusa Dias de. Equivalências: Do serviço de referência convencional a novos ambientes de redes digitais em bibliotecas. **Revista Brasileira de Biblioteconomia e Documentação**, São Paulo: Nova Série, v.1, n.1, p. 38-72, 1999.

MOOERS, Calvin N. Zatacoding applied to mechanical organization of knowledge. **American Documentation**, v.2, p. 20-32, 1951.

MORALES, Carmem Yasmina López. Infraestrutura y servicios de información en la biblioteca virtual. **Ciências de la Información**, v. 26, n. 3, sept., p. 115-122. 1995.

MUDDAMALLE, Manikya Rao. Natural language versus controlled vocabulary in information retrieval: a case study in soil mechanics. **Journal of the American Society for Information Science**, v.49, n. 10, p. 881-887, oct. 1998.

MUELLER, S. P. M. A ciência, o sistema de comunicação científica e a literatura científica. In: CAMPELLO, Bernadete Santos; CENDÓN, Beatriz Valadares; KREMER, Jeannette Marguerite (Org.). **Fontes de Informação para Pesquisadores e Profissionais**. Belo Horizonte, Ed. UFMG, p. 21-34, 2000.

NOGUEIRA, Luís. **Information retrieval**: uma abordagem baseada em agentes. Departamento de Engenharia Informática. Instituto Superior de Engenharia do Porto, Porto, jul. 2000.

OHIRA, Maria Lourdes Blatt; PRADO, Noêmia Schoffen. Bibliotecas virtuais e digitais: análise de artigos de periódicos brasileiros (1995/2000). **Ciência da informação**, Brasília, v.31, n.1, p.61-74, jan./abr. 2002.



ORTIZ-REPISO JIMÉNEZ, V. Qué enseñamos después Del MARC?. **Organización del Conocimiento en Sistemas de Información y Documentación**. Zaragoza, v.3, p. 217-225, 1999. Actas del III Encuentro de ISKO-España Getafe, 19 al 21 de noviembre de 1997.

PENNANEN, Mikko; VAKKARI, Pertti. Students' conceptual struture, search process, and outcome while preparing a research proposal: a longitudinal case study. **Journal of the American Society for Information Science and Technology**, v.54, n.8, p. 759-770, 2003.

PEREZ, Hélio Alvaro de Mello. **Modelagem de metadados para suporte a extração de dados em sistemas de informações ambientais**. Dissertação (mestrado) – Instituto militar de Engenharia, Rio de Janeiro, fev. 2000.

PÉREZ GUTIÉRREZ, Mario. El lenguaje de interrogación: una gramática formal para la recuperación de información. **Revista Española de Documentación Científica**, Madrid, v.23, n.3, p. 247-266, jul./sept. 2000.

PITASSI, Cláudio; LEITÃO, Sérgio Proença. Tecnologia de informação e mudança: uma abordagem crítica. **Revista de administração de empresas**, São Paulo, v.42, n.2, p.77-87, abr./jun. 2002.

QIN, J. Representation and Organization of Information in the Web Space: From MARC to XML. **Informing Science**, [S. l.], v.3, n.2, p. 83-87, 2000.

RIBEIRO, Célia Maria; SANTOS, Gildenir Carolino. **Acrônimos siglas e termos técnicos: arquivística, biblioteconomia, documentação, informática**. Campinas: Átomo, 2003.

RIBEIRO JÚNIOR, Divino Ignácio. Modelos de Recuperação de Informação. Recuperação de Informação – 5a. Fase. **Slides**. 2003. <<http://www.faed.udesc.br/cursobiblio/>>. Acesso em: 25 mar. 2003.

RIBEIRO JÚNIOR, Divino Ignácio. **Notas de aula**. 2003. Disponível em: <<http://www.faed.udesc.br/cursobiblio/>>. Acesso em: 25 mar. 2003.

ROBINS, David. Interactive Information Retrieval:Context and Basic Notions. **Informing Science: Special Issue on Information Science Research**, v.3, n.2, p. 57-61, 2000.

ROBREDO, Jaime, **Da ciência da informação revisitada aos sistemas humanos de informação**. Brasília: Thesaurus; SSRR Informações, 2003. 262 p.

RODRÍGUEZ BRAVO, Blanca; SANTOS de PAZ, Lourdes. Del documento digital a la biblioteca virtual. **Scire: representación y organización del conocimiento**. Universidad Zaragoza, v.8, n.2, p. 43-52, jul./dic. 2002.

ROWLEY, Jennifer E. **Abstracting and indexing**. 2.ed., London: Clive Bingley, 1988.

ROWLEY, Jennifer E. **A biblioteca eletrônica: 2ª edição de informática para bibliotecas**. Brasília: Brinquet de Lemos/Livros, 2002, 399 p.

SALVADOR OLIVÁN, José Antonio; ANGÓS ULLATE, José Maria; MARIA JESÚS, Fernández Ruiz. Metadatos en los documentos HTML: una ayuda para la recuperación de información. **Scire: representación y organización del conocimiento**. Universidad Zaragoza, v.5, n.2, p. 47-61, jul./dic. 1999.

SANT'ANA, Ricardo César Gonçalves; SANTOS, Plácida L. V. A. Costa. Transferência de informação: análise de fatores para identificação do valor de unidades de conhecimento registrado. In: VIDOTTI, Ap. B. Gregório (coord). **Tecnologia e conteúdos informacionais: abordagens teóricas e práticas**. São Paulo: Polis, 2004. p. 53-75.

SARACEVIC, Tefko et al. Nature of the interaction between users and intermediaries in on-line searching: a qualitative analysis. **Proceedings of the Annual Meeting of the American Society for Information Science**. v. 27, p. 47-54, 1990.

SARACEVIC, Tefko. Ciência da informação: origem, evolução e relações. **Perspectivas em ciência da informação**. Belo Horizonte, v.1, n.1, p. 41-62, jan./jun. 1996.

SARACEVIC, Tefko. Modeling interaction in information retrieval (IR): a review and proposal. In: Hardin, S., **Annual Meeting of the American Society for Information Science**. 59ed. Silver Spring, MD: American Society for Information Science, p. 03-09, 1996.

SHAPIRO, Carl ; VARIAN, Hal R. **A economia da informação**: como os princípios econômicos se aplicam a era da Internet. 5ed., Rio de Janeiro: Campus, 1999. Tradução de: Information rules por Ricardo Inofosa. 397 p.

SILVEIRA, Maria de Lourdes. **Recuperação da informação utilizado tesauros**: um estudo de caso na área jurídica. 2002. Tese de doutorado. UFMG. 2002. Disponível em: <http://www.dcc.ufmg.br/pos/html/spg2002/anais/maria/maria.html>. Acesso em: 15 maio 2003.

SOERGEL, D. Indexing and retrieval performance: the logical evidence. **Journal of ASIS**. v. 45, n.8, p. 589-599, 1994.

SOUZA, M. I. F. ; VENDRUSCULO, L. G. ; MELO, G. C. Metadados para a descrição de recursos de informação eletrônica : utilização do padrão Dublin Core. **Ciência da Informação**, Brasília, v. 29, n.1, p. 93-102, jan./ abr. 2000.

SPARCK-JONES, K.; WILLET, P. (editors). **Readings in Information Retrieval**. California: Morgan Kaufmann Publishers, Inc., 1997.

SPINK, Amanda. Study of interactive feedback during mediated information retrieval. **Journal of the American Society for Information Science**, v.48, n.5, p. 382-394, 1997.

SPINK, Amanda; SARACEVIC, Tefko. Interaction in Information Retrieval: Selection and Effectiveness of Search Terms. **Journal of the American Society for Information Science**, v.48, n.8, p. 741-761, 1997.

SPINK, A.; GOODRUM, A.; ROBINS, D. Elicitation behavior during mediated information retrieval. **Information Processing & Management**. v. 34, n. 2/3, p. 257-273, 1998.

SPINK, Amanda. Toward a theoretical framework for information science. **Informing Science**. v.3, n.2, 2000.

STIX, Gary. Tempo Real. **Scientific American: Brasil**, n.5, p.50-53, out. 2002.

STRZALKOWSKI, TOMEK (editor). **Natural Language Information Retrieval**. Kluwer Academic Publishers, 1999. 385 p.

STURGES, Paul. Gatekeepers and other intermediaries. **Aslib Proceedings**. v. 53, n.2, p. 62-67, Feb. 2001.

TAKAHASHI, Tadao (org). Ministério da Ciência e Tecnologia. **Sociedade da informação no Brasil**: livro verde. Brasília, DF, 2000. 195p.

TARAPANOFF, Kira et al. Inteligência obtida pela aplicação de data mining em base de teses francesas sobre o Brasil. **Ciência da informação**, Brasília, v.30, n.2, p.20-28, maio/ago. 2001.

TASSO, Carlo; VENUTI, Fabio; BRAJNIK, Giorgio; MIZZARO, Stefano. Strategic help in user interfaces for information retrieval. **Journal of the American Society for Information Science and Technology**. v.53, n.5, p. 343-358, mar. 2002.

TENOPIR, Carol. Eletronic publishing: research issues for academic librarians and users. **Library Trends**, v. 51, n. 4, p. 614-635, 2003.

TRAMULLAS SAZ, Jesús. Perspectivas en recuperación y explotación de información eletrónica: el “data mining”. **Scire**: represnetación y organización del conocimiento. Universidad Zaragoza, v.3, n.2, p. 75-83, jul./dic. 1997.

VALENTIM, Lígia Pomim. Atuação e perspectivas profissionais para o profissional da informação. In: VALENTIM, Lígia Pomim (Org.) **O profissional da informação**: formação, perfil e atuação profissional. São Paulo: Polis, 2000. cap. 7, p. 135-152.

VIEIRA, Sonia. **Elementos de estatística**. 3.ed. São Paulo: Atlas, 1999.

WILSON, T. D. Models in information behaviour research. **Journal of Documentation**. London, v.55, n.3, p. 249-270, june 1999.